

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:29:52 ; Search time 18 Seconds

(without alignments)
711.054 Million cell updates/sec

Title: US-09-913-414-5

Perfect score: 2165

Sequence: 1 MKPPKIQRPDMFOAVARIG.....VLQYLFSPRLRLRTMREE 435

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued_Patents_AA.*

2: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*

3: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*

4: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*

5: /cgn2_6/ptodata/1/1aa/6B.COMB.pep.*

6: /cgn2_6/ptodata/1/1aa/6C.COMB.pep.*

7: /cgn2_6/ptodata/1/1aa/6D.COMB.pep.*

8: /cgn2_6/ptodata/1/1aa/6E.COMB.pep.*

9: /cgn2_6/ptodata/1/1aa/6F.COMB.pep.*

10: /cgn2_6/ptodata/1/1aa/6G.COMB.pep.*

11: /cgn2_6/ptodata/1/1aa/6H.COMB.pep.*

12: /cgn2_6/ptodata/1/1aa/6I.COMB.pep.*

13: /cgn2_6/ptodata/1/1aa/6J.COMB.pep.*

14: /cgn2_6/ptodata/1/1aa/6K.COMB.pep.*

15: /cgn2_6/ptodata/1/1aa/6L.COMB.pep.*

16: /cgn2_6/ptodata/1/1aa/6M.COMB.pep.*

17: /cgn2_6/ptodata/1/1aa/6N.COMB.pep.*

18: /cgn2_6/ptodata/1/1aa/6O.COMB.pep.*

19: /cgn2_6/ptodata/1/1aa/6P.COMB.pep.*

20: /cgn2_6/ptodata/1/1aa/6Q.COMB.pep.*

21: /cgn2_6/ptodata/1/1aa/6R.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	521.5	24.1	443	2	US-09-005-232A-3
2	510.5	23.6	443	2	US-08-620-605D-3
3	313.5	14.5	477	3	US-08-772-270A-13
4	262	12.1	478	3	US-08-772-270A-5
5	201.5	9.3	425	2	US-08-853-659A-38
6	118.5	5.5	796	4	US-08-868-699A-2
7	118.5	5.5	796	4	US-09-757-014-2
8	115	5.3	1090	4	US-09-085-199B-5
9	113.5	5.2	457	4	US-08-924-629C-6
10	112	5.2	531	4	US-09-134-001C-4920
11	112	5.2	914	4	US-09-085-199B-4
12	109	5.0	1664	1	US-09-599-652-2
13	109	5.0	1664	2	US-08-642-846-2
14	109	5.0	1664	2	US-09-264-604-2
15	107	4.9	747	2	US-08-816-693A-51
16	107	4.9	747	2	US-08-885-291-51
17	107	4.9	747	4	US-09-496-672-51
18	106.5	4.9	591	1	US-08-178-738-5
19	106.5	4.9	591	2	US-08-628-145-5
20	106	4.9	2285	4	US-09-308-375-2
21	105	4.8	693	4	US-08-235-836C-68
22	104	4.8	728	4	US-09-134-001C-4968
23	104	4.8	1146	4	US-08-914-999-6
24	103	4.8	379	1	US-08-279-270A-1
25	102.5	4.7	529	4	US-08-887-534A-74
26	102.5	4.7	635	4	US-09-341-833A-7
27	102.5	4.7	677	4	US-09-341-833A-8

28	102.5	4.7	1048	4	US-08-887-534A-85	Sequence 85, Appl
29	102.5	4.7	1780	1	US-08-769-309A-5	Sequence 5, Appl
30	102.5	4.7	1780	3	US-08-994-570-5	Sequence 5, Appl
31	102	4.7	659	4	US-08-894-818B-1	Sequence 1, Appl
32	102	4.7	659	4	US-09-445-472-12	Sequence 12, Appl
33	102	4.7	756	4	US-09-085-199B-9	Sequence 9, Appl
34	102	4.7	896	1	US-08-095-737-2	Sequence 2, Appl
35	102	4.7	896	1	US-08-480-145-2	Sequence 2, Appl
36	102	4.7	896	2	US-08-477-389-2	Sequence 2, Appl
37	101.5	4.7	534	4	US-09-103-664A-2	Sequence 2, Appl
38	101.5	4.7	693	4	US-08-235-836C-72	Sequence 72, Appl
39	101	4.7	635	4	US-09-341-833A-9	Sequence 9, Appl
40	101	4.7	1388	2	US-08-685-576-4	Sequence 4, Appl
41	101	4.7	2101	1	US-08-195-487-4	Sequence 4, Appl
42	101	4.7	2101	5	PCT-US93-06160-4	Sequence 4, Appl
43	100.5	4.6	635	4	US-09-341-833A-6	Sequence 6, Appl
44	100	4.6	1388	2	US-08-685-576-1	Sequence 1, Appl
45	99.5	4.6	955	1	US-08-006-676B-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-005-232A-3
Sequence 3, Application US/09005232A
Patent No. 5922568
GENERAL INFORMATION:
APPLICANT: SHIBATANI, TAKEJI
APPLICANT: AKATSUKA, HIROYUKI
TITLE OF INVENTION: GENE PARTICIPATING IN THE MECHANISM OF
TITLE OF INVENTION: SECRETION OF ESTERASE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
STREET: PO BOX 747
CITY: FALLS CHURCH
STATE: VA
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/005,232A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: SVENSSON, LEONARD R
REGISTRATION NUMBER: 30,330
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)-205-8000
TELEFAX: (703)-205-8050
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 443 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-005-232A-3
Query Match
Best Local Similarity 29.3%, Pred. No. 4, 1e-41;
Matches 133; Conservative 88; Mismatches 194; Indels 39; Gaps 7;
OY 6 IORPTDNF-----QAVARIGIIALTVGLGMAFAPLDSAVTANGVSAEVS-Q 56
DB 5 IGEPODSTTEIPDERFRFTMGMLVVGIGLGFPLAAMAFAPLDKGVASPGSVTVSCNRK 64

TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 13:

ADDRESS: Intellectual Property Services
ADDRESS: Battelle Memorial Institute
ADDRESS: PNNL P.O. Box 999
STREET: Washington Way

```

: CITY: Richland
: STATE: Washington
: COUNTRY: U.S.A.
: ZIP: 99352
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette 3.50", 1.44 Mb storage
: COMPUTER: IBM PC/XT/AT
: OPERATING SYSTEM: MS-DOS
: SOFTWARE: Word Processor (WordPerfect 5.1)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/853,659A
: FILING DATE: Unknown
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: FILING DATE: n/a
: INFORMATION FOR SEQ ID NO: 38:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 425 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
US-08-853-659A-38

```

```

Query Match          9.3%; Score 201.5; DB 2; Length 425;
Best Local Similarity 21.0%; Pred. No. 7.1e-11;
Matches 94; Conservative 93; Mismatches 187; Indels 73; Gaps 19;

```

```

QY 7 QRPDNEQAVARIGYIATLFGVGLGMAFAPLDSAVIANGVSA-EVSODVHLEGGM 65
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3 RRGSDHLMH-----IISLTL-IILTYFELINSVHGGVITTKMADLISLKGST 55
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 66 LAKLVNREGVYKAGQV---FELDPTQANAAGITRNQYVALKAMEARLLAERDQPSI 122
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 56 IQDIYVNEGDPVKKGELLAKVNL-----LQKEYORYTQKGYL--DKDVN-EI 102
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 123 SEPADLSORADPMV--ARATADQAOETERRQTIQGOVDLMNQRLOYOSEIGSDMQT 180
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 103 SFILDKENESGLITLDGTRLSNKEVK--ANIELVHSQIRAKELKTSLSDEISGLQEL 160
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 181 QGLKDOLFIDEIDLKLYDKGLV-----RPRLALEARAGSLSGISGRITADRS 233
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 161 SSKKEELALLAEINILSPLVKKGISPTYNFLNKKOATIKVKSLENDLESTI-TLKRD 219
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 234 KAVGASDTQLKVKROIKEFEQVQSITETRV--RLAEVTEKEVVASDAOKRIKIVSPV 291
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 220 ELV--VNDIEALNLELRLSLKISKNLOELVNVSTLKVEKIQNEED-----IYSPV 271
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 292 NGTAQ--NLRFETGAVYVRAEPLVDIAPDEAFVIAHQPTQVDVNHMGMTVEVRLPA 349
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 272 DGVLYIKKSAITTHGVIQADLLFEIKPKVYTMADVKIILPKYRDQIYVDEAVKLDVQS 331
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 350 FHSAGNDPERHDPVAVADRISDPQK-----ARLEGIYRVYDKV-----LPHLNG 397
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 332 I-----IQPKIKSYNATIDNIDSPSYENGTGIIQRYKVIILADVNNDDLRLWLPK 382
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 398 RVTAGMPAQVIVPTGERTVLOYLFSPL 424
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 383 ----GMTVDASVITGKHSIMEYLLSPL 405
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 6
US-08-868-699A-2
: Sequence 2, Application US/08868699A
: Patent No. 6204019
: GENERAL INFORMATION:
: APPLICANT: O'Dwyer, Karen
: APPLICANT: Perry, Caroline
: APPLICANT: Warren, Richard L.
: TITLE OF INVENTION: No. 6204019e1 Compounds
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dechert, Price & Rhoads
: STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre

```

```

: CITY: Philadelphia
: STATE: PA
: COUNTRY: USA
: ZIP: 19103-2793
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: Windows
: SOFTWARE: FastSeq for Windows Version 2.0b
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/868,699A
: FILING DATE: 04-JUN-1997
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Falk, Stephen T
: REGISTRATION NUMBER: 36,795
: REFERENCE/DOCKET NUMBER: GM10012
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 215-994-2488
: TELEFAX: 215-994-2222
: TELEX:
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 796 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
US-08-868-699A-2

```

```

Query Match          5.5%; Score 118.5; DB 4; Length 796;
Best Local Similarity 20.6%; Pred. No. 0.013;
Matches 63; Conservative 52; Mismatches 120; Indels 71; Gaps 11;

```

```

QY 10 TDNQAVARIGYIATLFGVGLGMAFAPLDSAVIANGVSAEVSQVQHL----- 61
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 358 TTFQNLFLKLESPSGMTATGCKLGESEFFDLYSKI---VQAPTDAIQRIDEPDKVFR 413
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 62 -----EGGMLAKILVREGKVK--AGQVLEFELDPPTQANAAGITRNQYV 103
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 414 SVDEKNIAIMHDIYELHETGRPVLLITRTAEAEYFSEVLFPMD-----IPNLLI 464
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 104 ALK-AMEARLLAERDQPSISFPADLTSORADPMVARAIDQAOETERRQTIQGOVDLM 162
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 465 AQVAKAEQOMIAEAGQIGSMVTATSMAGRTDIKIGCV-----EALGGLAVII 513
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 163 NAQRLQYQSEIEGIDROTQGLKQDLG-----FIEDELIDLKLYDKGLVRRPLLA 213
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 514 HE-----HMENSRVDRQLRGHSGRGDGGSSCIYISLDVLY--KRMSDSNLAENNOLYS 566
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 214 LEARAGSLSGISGRITLDRSKAVOGASDQO-LKVRQIKOFEFEQVS---QSTFETVRRLA 269
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 567 LDAORLQSQNLFNKRVKQIYVKAQRISEBQGVKAREANFEKISIQRLDLYVERNRVL 626
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 270 EVTEKE 275
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 627 EIDDAE 632
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 7
US-09-757-014-2
: Sequence 2, Application US/09757014
: Patent No. 6348342
: GENERAL INFORMATION:
: APPLICANT: O'Dwyer, Karen
: APPLICANT: Perry, Caroline
: APPLICANT: Warren, Richard L.
: TITLE OF INVENTION: No. 6348342e1 Compounds
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dechert, Price & Rhoads

```

```

: CITY: Philadelphia
: STATE: PA
: COUNTRY: USA
: ZIP: 19103-2793
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: Windows
: SOFTWARE: FastSeq for Windows Version 2.0b
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/868,699A
: FILING DATE: 04-JUN-1997
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Falk, Stephen T
: REGISTRATION NUMBER: 36,795
: REFERENCE/DOCKET NUMBER: GM10012
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 215-994-2488
: TELEFAX: 215-994-2222
: TELEX:
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 796 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
US-08-868-699A-2

```

```

STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103-2793
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/757,014
FILING DATE: 09-Jan-2001
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/868,699
FILING DATE: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Falk, Stephen T
REGISTRATION NUMBER: 36,795
REFERENCE/DOCKET NUMBER: GM10012
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-994-2488
TELEFAX: 215-994-2222
TELEX: <unknown>
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 796 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-757-014-2

Query Match      5.5%; Score 118.5; DB 4; Length 796;
Best Local Similarity 20.6%; Pred. No. 0.013;
Matches 63; Conservative 52; Mismatches 120; Indels 71; Gaps 11;

OY 10 TDNFOAVARIGGIALFVGLGMAAFAPLDSAVIANGVSAEVSODVQH----- 61
DB 358 TITFONLKRLEFSFGMTATGIGSEFPDLYSKI-----VQAPIDKAIQRIDEPKYFR 413
OY 62 -----EGMKLAKILVREGKVK-AGGVLELDEPTQANAAAGITRNOYV 103
DB 414 SVDEKNIAMIHDIHELHETGRPVLLITRTAEAEYSEVLFGMD-----IPNNLLI 464
OY 104 ALK-AMEARLLAEQRORSISFPADLTQSORADPMVARRAIADQAFTERQTIQGVDM 162
DB 465 AONVAKEAQMIAEAGQISMTVATSMAGRTDIKLGEGV-----EALGGLAVI 513
OY 163 NAORLOVSEIEGIDROGKLDQGL-----FIEDELIDRLKLYDKGLVPRPRLIA 213
DB 514 HE-----HMENSRVDRQLRGSRGQDPSGSCITISLDDYLV--KMSDSNLAENNOQLYS 566
OY 214 LEARAGSLSGISGLTLADRSKAVGASDTQ-LKVRQIKOEFEQVS---QSITETRVRLA 269
DB 567 LDAQRLSQSNLFRKVKQIVVAKRISSEQGVKAREMANEFKKSISIQDLYEERNRVL 626
OY 270 EYTEKE 275
DB 627 EIDDAE 632

```

RESULT 8
US-09-085-199B-5
Sequence 5, Application US/09085199B
Patent No. 6235879

GENERAL INFORMATION:
APPLICANT: Hayden, Michael R.
APPLICANT: Hacken, Abigail
APPLICANT: Hug, A.H.M. Mahabul
APPLICANT: Chopra, Vikramjit Singh
APPLICANT: Kalchman, Michael

```

TITLE OF INVENTION: Apoptosis Modulators That Interact with the
TITLE OF INVENTION: Huntington's Disease Gene
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESS: Opedahl & Larson
STREET: PO Box 5270
CITY: Frisco
STATE: CO
COUNTRY: USA
ZIP: 80443-5270
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Kb storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS DOS 5.0
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/085,199B
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Larson, Marina T.
REGISTRATION NUMBER: 32038
REFERENCE/DOCKET NUMBER: USC.P-013052
TELECOMMUNICATION INFORMATION:
TELEPHONE: (970) 668-2050
TELEFAX: (970) 668-2052
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1090
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: no
ORIGINAL SOURCE:
ORGANISM: human
FEATURE:
OTHER INFORMATION: Huntington-Interacting protein
US-09-085-199B-5

```

Query Match 5.3%; Score 115; DB 4; Length 1090;
Best Local Similarity 20.1%; Pred. No. 0.045;
Matches 82; Conservative 72; Mismatches 131; Indels 122; Gaps 20;

```

OY 47 NGVSAEVSODVQHE---GMLAKILVREGKVKAGVLELDEPTQANAAAGITRNOYV 103
DB 419 NGVNKDEKDHILRLRLREISGLKQL---ENKKTESQRYVLQKHVSELEADLAEQHL 475
OY 104 ALKAME--ARLLAEQRORSISFPADLTQSORADPMVARRAIADQAFTERR----- 152
DB 476 RQQAADDEFLRAELD-----ELRRQREDTEKAKRSISEIRKKAQANQORYSKLKE 526
OY 153 ---QTIQGVDM--NAQ-----RLOYSEIEGIDROGKLDQGLFED-----ELI 195
DB 527 KYSELVQNHADILRRNAEYTKOVSMARQAQVD-LEREKKELEDLSERISDQGRTOFOL 585
OY 196 DLRLKYDGLVPRPRLIALEARAGSLGS-----IGRLTPADRSKAVGASDTOLK 245
DB 566 EYLESIKQELATSOR--ELQVLYGSLSTSAQSEANMAAFELKEKRSIVSGAHHREE 643
OY 246 VROIKOEFEQVSOSITETRVRLAEVTEKEVVASDAQRIKIVSPVNGTAONLRFETGCA 305
DB 644 LSAIRKE-----LQDTQLKLASTEESMCQLAKDQKMLLV---GSRKAAEQVIODA 691
OY 306 VVRAAE-PLVYDA-----PRD-----EAFVIOAHF 329
DB 692 LQLEEPPLISGASADHLSTVTSISSCIEOLEKXSMQVYACPPDISGLLHSTITLHL 751
OY 330 QPTDVNVMGMVTEVRLEPAFHSAGNPDPERHDPVAVADRISDPQKQ 376
DB 752 TS---DATAHGATITCLRAP-----PEP-----ADSLTEACKQ 780

```

RESULT 9

```

US-08-924-629C-6
; Sequence 6, Application US/08924629C
; Patent No. 6403082
; GENERAL INFORMATION:
; APPLICANT: Stiles, Michael E.
; APPLICANT: Vedernas, John C.
; APPLICANT: van Belkum, Marius J.
; APPLICANT: Morobo, Randy W.
; APPLICANT: Morobo, Rodney J.
; APPLICANT: Greer, G. Gordon
; APPLICANT: McMullen, Lynn M.
; APPLICANT: Leisner, Jorgen J.
; APPLICANT: Poon, Alston
; APPLICANT: Franz, Charles M.A.P.
; TITLE OF INVENTION: No. 6403082elBacteriocins, Transport and Vector System and Method
; FILE REFERENCE: 660.000505
; CURRENT APPLICATION NUMBER: US/08/924,629C
; CURRENT FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 60/026,257
; PRIOR FILING DATE: 1996-09-05
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 457
; TYPE: PRT
; ORGANISM: lcad;
US-08-924-629C-6

```

```

Query Match          5.2%; Score 113.5; DB 4; Length 457;
Best Local Similarity 18.4%; Pred. No. 0.017;
Matches 85; Conservative 76; Mismatches 195; Indels 105; Gaps 16;

QY 7 QRPDNEQAVARIGYIATLFTVGLGMAFAPLDSAVIANG-VYSAEVSODVOHLEGGM 65
   || || || || || || || || || || || || || || || || || || || ||
DB 14 QRRRNEFTLIIV--IFLLVFIILFLFAKREIYVAKGELIIPAVLSDIGSTNNA 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 66 LAKLIVREGYVAKGQVLELDPTQANAGITRNOYVALKAMEARLAEEDQRPST--- 122
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 71 IDSNQLETKYVKKGDLVFTSGNEKISQLTQO--LNNLNDRLKSLDLYKOSIYNG 127
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 123 -----SEPAD-----LTSQRADPMVARAIADBOAQFTRRRTQIGOVNDLM 162
   || || || || || || || || || || || || || || || || || || || ||
DB 128 RSEGGTQFGYDLSFNQYMAQVDTLISEFNOOSSDKOTADQQAHH-----QIVL 178
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 163 -----NAORLYQOSEIEGIDROTQGLKDLGFIED---ELIDRLKLYDKGLVPRRL 211
   || || || || || || || || || || || || || || || || || || || ||
DB 179 KOGSKNNQOALANYQAILTSINSNTKPTNNPYQAIYDNYSAOLKSAQTTDDKQYKQVAL 238
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 212 L-----ALBARAGSLSGISGRLADRSKAVQASDPTQLVROKQEFFEQVOSITET 264
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 239 SNVOQOIQDQLTSSSYDSQIAGIT--KSGPLSQSSTLIDKADKQOOLASAKKEINDQ 295
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 265 RVRLAEVTEREVASDAQRIKIVSPVNGTAQ-----NLRFTEGAVVRAAEPLVDIAP 318
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 296 QOSIDELKAKASSANEDQDVIVAKPEDGILHLATDKIKIYFPGTITAIQYRKL--T 352
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 319 EDEAFVIOAHFOPTDY-----DNVHGMVTEVRLPAFHSAGNPDEPRHDV 364
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 353 QKTALNVEYVPAASNIILKQROAIRFVANOVTKPLTLNGLTISISSA-----PI 403
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 365 AVADRISPOKQARLF-----LGIVRDVYKOLPRLRGRTV 400
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 404 A-----SKEGSFYKLAVTIOASKIDREQIKYGLNGRIT 436
   : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 10
US-09-134-001C-4920
; Sequence 4920, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

```

```

; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4920
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4920

```

```

Query Match          5.2%; Score 112; DB 4; Length 531;
Best Local Similarity 18.7%; Pred. No. 0.029;
Matches 69; Conservative 67; Mismatches 153; Indels 80; Gaps 11;

QY 44 VIANGVYSAEVSODVOH--LEGMLAKILY---REGKYKAGQVLELDPTQ--ANAA 94
   : : || || || || || || || || || || || || || || || || || || ||
DB 25 IILGVVGYIYARNLILHOKQVQARQTADDIVSYANKKADNKKKELLEAKENQILKEQA 84
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 95 AGITRNOYVALKAMEARLAEEDQRPSTSPADLTSQRADPMVARAIADBOAQFTERRQT 154
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 85 ENELRERGERELQROETRLQKE--NLDRKSDLLDKRDE---ILEQESKLEERQOQ 136
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 155 IGGQVDMNAORLYQOSEIEGIDROTQ--GLKDQGLFIEDELIDRLKLYDKGLVPRRL 212
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 137 VDAKESSVQTLIMHLELEERISGLTOBEAVKBELOVVEEL----- 178
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 213 ALEARAGSLSGISGRLADRSKAVQASDPTQLVROKQEFFEQVOSITETRVRLAEVT 272
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 179 -----SODIALILVKEKEKEKEKEDVTAKEL---LATTV 209
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 273 EKEVYASDAQRIKIVSPVNGTAQNLRFTEGAVVRAEPL-----VDIAPDEAVIO 326
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 210 QRLAEHTTESTVSVNLPDEMKGRIIGREGNIRITLTLTGILTIIDTPE--AVIL 266
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 327 AHFOPTDVNVMQVTEVRLPAFHSAGNPDEPRHDVAVADRISPOKQARLFGLIVRV 386
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 267 SCFPIREIRLARTALVNLVSGRHRHPGRIEDMVAKARKEVDIIRDAEGEATF-----EI 321
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 387 DYKOLPRL 395
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 322 NVNHHMPDL 330
   : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 11
US-09-085-199B-4
; Sequence 4, Application US/09085199B
; Patent No. 6235879
; GENERAL INFORMATION:
; APPLICANT: Hayden, Michael R.
; APPLICANT: Hackam, Abigail
; APPLICANT: Hug, A.H.M. Mahbulul
; APPLICANT: Chopra, Vikramjit Singh
; TITLE OF INVENTION: Apoptosis Modulators That Interact with the
; TITLE OF INVENTION: Huntington's Disease Gene
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oppe Dahl & Larson
; STREET: PO Box 5270
; CITY: Frisco
; STATE: CO
; COUNTRY: USA
; ZIP: 80443-5270
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Kb storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS DOS 5.0
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:

```



```

: APPLICATION NUMBER: US/09/085,199B
: FILING DATE:
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Larson, Marina T.
: REGISTRATION NUMBER: 32038
: REFERENCE/DOCKET NUMBER: UDC.P-013052
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (970) 668-2050
: TELEFAX: (970) 668-2052
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 914
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: HYPOTHEICAL: no
: ORIGINAL SOURCE:
: ORGANISM: human
: FEATURE: Huntington-interacting protein
: US-09-085-199B-4

```

```

Query Match          5.2%; Score 112; DB 4; Length 914;
Best Local Similarity 20.1%; Pred. NO. 0.067;
Matches 82; Conservative 72; Mismatches 131; Indels 122; Gaps 20;

```

```

QY 47 NGVSAEYSDVOHLE---GGMALIKILREGEKAVAGOVLELDPDPTQANAAAGITRNOYV 103
DB 243 NGVNDEKDHLEIRLYRELSGAKOL---ENKKTESQRYVILQKHVSELEADLAEOQHL 299
QY 104 ALKAME--ARLIAERDQRPSSIFPADLTSORADPVAVARAIDE---QAQETERR----- 152
DB 300 RQQAADDEFLRAELD-----ELRRQREDTEKAQRSISEIRKQANQRYSKLKE 350
QY 153 ---QTIQGVNDL--NAQ-----RLQYQSEIEGIDRQTOGLKQDGFIED-----ELI 195
DB 331 KYSELVQNHADLKRKNAEYTVQVSMARQAYD-LEKEKKELEDSLEKISDQCRKTOEQL 409
QY 196 DLRLKLYDGLVPRRLALLEARAGSLGS-----IGRLTADRSKAVAGASDTOLK 245
DB 410 EYLESKQELGTSOR--ELQVLOGSLSTSAQSEANMAAFELKEKERSLVSGAHHREE 467
QY 246 VRQIQEPEEYQSITETRVRLAEVTEKEVVASDAQRIKIKIVSVNGTAQNLRFETEGA 305
DB 468 LSAIRKE-----LQDTQLKLAESTEESMQLAKDKRMILY-----GSRKAAEQVIDA 515
QY 306 VYRAAE-PLVYDA-----PRD-----EAFVIOAHF 329
DB 516 LNLDEEPLISCAGSADHLSTVTSISSCIEQLKESMSQYLACPDIDISGLLHSTITLHL 575
QY 330 QPTVDYNHMGVTEVRLPAFHSAGNPDPERHDPVAVADRISSDPQKQ 376
DB 576 TS---DATAHGATTCRLAP-----PEP-----ADSLTEACKQ 604

```

```

RESULT 12
US-09-599-652-2
: Sequence 2, Application US/09599652
: Patent No. RE37741
: GENERAL INFORMATION:
: APPLICANT: HOSTETTER, MARGARET K.
: APPLICANT: GALE, CHERYL A.
: APPLICANT: BENDEL, CATHERINE M.
: APPLICANT: TAO, NIAN-JUN
: APPLICANT: KENDRICK, KATHLEEN
: TITLE OF INVENTION: CANDIDA ALBICANS GENE, INTEGRIN-LIKE
: TITLE OF INVENTION: PROTEIN, ANTIBODIES, AND METHODS OF USE
: NUMBER OF SEQUENCES: 12
: CORRESPONDENCE ADDRESS:
: ADDRESS: MUEITING, RAASCH, GRABARDT & SCHNAPPACH, P.A.
: STREET: 119 NORTH FOURTH STREET, SUITE 203
: CITY: MINNEAPOLIS
: STATE: MINNESOTA

```

```

: COUNTRY: USA
: ZIP: 55401
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/599,652
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/642,846
: FILING DATE: 03-MAY-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: MUEITING, ANN M.
: REGISTRATION NUMBER: 33,977
: REFERENCE/DOCKET NUMBER: 110.00280101
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 612-305-1217
: TELEFAX: 612-305-1228
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1664 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-09-599-652-2

```

```

Query Match          5.0%; Score 109; DB 1; Length 1664;
Best Local Similarity 19.5%; Pred. NO. 0.32;
Matches 80; Conservative 61; Mismatches 124; Indels 146; Gaps 20;

```

```

QY 92 NAAAGITRNOYALAMERARLIAERDQRPSSIFPADLTSORADPVAVARAIDEQAQETERR 151
DB 974 NAKKVTODEYI---NAKLVDQPKKNSI-----VTPDEDEYEEL 1010
QY 152 ROT-----IQGVNDLMAOIRLQYQSEIEGIDRQTOGLKQDGFIEDELIDLRK- 199
DB 1011 QOTASIHNAITDSSLYGRPDSISTDMLRLSD-ELKRPRTALLSADRLFMEVEVRLPSN 1069
QY 200 -----LYDKGLVPRRLALLEARAGSLGSIGRTLADRSKAVAGASDTOLKVRQI 249
DB 1070 SVLVHPGAGAATNSMLDEPDELINSPARVNSNNDV-----AISGNAST-ISPQNL 1122
QY 250 KOEFPEQ--VQSITET-----TRVRLAEVTEKEVVASDAQ 282
DB 1123 DMNFDQATIGQIKQIEOPARSKANTVRGDDDLASAPETPRTPRKESISSKPARLSAS 1182
QY 283 KR---IKIVSPVNGTAQNLRFETEGAVYRAEPLVYDAPEDEAFVIOAHFOPTVDVN--V 337
DB 1183 PRKSPKIGSPV-----RVTKKNGSIAGIEPIPKATHKP-----KSPQGNELSNKV 1230
QY 338 HMGVTE-----VRLPAFHSAGN---PDP-----BRHD-- 362
DB 1231 RDGGISPSGSEHQOHNPSMWSVPSQYTDATSTVDENKMDYQHKPREKOKKNNHNNHNN 1290
QY 363 -----PVAVDKISDP--QKQARLF---LGIYRVVDYKQRLRYLRGRT 400
DB 1291 HHKQKTDIPGVVDDEIPVGLQDERGKLFRRVLGINKINMLPDINTH-KGRFT 1340

```

```

RESULT 13
US-08-642-846-2
: Sequence 2, Application US/08642846
: Patent No. 5686151
: GENERAL INFORMATION:
: APPLICANT: HOSTETTER, MARGARET K.
: APPLICANT: GALE, CHERYL A.
: APPLICANT: BENDEL, CATHERINE M.
: APPLICANT: TAO, NIAN-JUN
: APPLICANT: KENDRICK, KATHLEEN

```

TITLE OF INVENTION: CANDIDA ALBICANS GENE, INTEGRIN-LIKE
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: MUEITING, RAASCH, GERHARDT & SCHWAPPACH, P. A.
STREET: 119 NORTH FOURTH STREET, SUITE 203
CITY: MINNEAPOLIS
STATE: MINNESOTA
COUNTRY: USA
ZIP: 55401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/642,846
FILING DATE: 03-MAY-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: MUEITING, ANN M.
REGISTRATION NUMBER: 33,977
REFERENCE/DOCKET NUMBER: 110.00280101
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-305-1217
TELEFAX: 612-305-1228
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1664 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-642-846-2

Query Match 5.0%; Score 109; DB 2; Length 1664;
Best Local Similarity 19.5%; Pred. No. 0.32;
Matches 80; Conservative 61; Mismatches 124; Indels 146; Gaps 20;

QY 92 NAAAGITRNOVALKAMARLARLAERDQPSISFPADLTSGRADPMVARAIADQOFTER 151
DB 974 NAKKGVTODEYI-----NAKLVDQPKKNSI-----VTDPEDRYEEL 1010

QY 152 ROT-----IOGVDMNNAORLOYOSEIEGIDRQGLKQDLGFIEDELIDLRK- 199
DB 1011 QOTASINMATIDSSYIGRPDISITDMLPYLSD-ELKKPTALLSADRLFMEOEVNRLRSN 1069

QY 200 -----LYDKGLVPRRLALLAERAGSLSGISGRLTADRKAAGVAGASDTOLKVRQI 249
DB 1070 SVLVHPGAGATNSMPEPDELINSPARVNSNSDNV-----AISGNAST-ISFNOL 1122

QY 250 KOEFFEQ--VSOSITE-----TVRLAEVTEKEVVASDAQ 282
DB 1123 DMNDDQATIQKIOEOPASKANTVRGDDGLASARETPRTPTKESISSKRAKLSAS 1182

QY 283 KR---IKIVSVNGTAQNLREFTEGAVVRAERLVDIAREDEAFVIAHFOPTVDN--V 337
DB 1183 PRKSPIKIGSV-----RVTKKNGSIAGIEPIPKATHKP-----KKSFGNGEISNKHV 1230

QY 338 HMGWTE-----VRLPAHSAGN---PDP-----ERND-- 362
DB 1231 RDGGISPSGSEHQHNPMSVSPSYQTDTATSTVDENKDVQYHPRKQOKHNNHNNHNN 1290

QY 363 -----PVAVADRISDP--OKQARLF--LGIIVRVQKOLRPHLGKRV 400
DB 1291 HHKQKTDIPGVVDEIPVQGLQERKILFFRVLGIKINIMLPDINTH-KGRFT 1340

RESULT 14
US-09-264-604-2
Sequence 2, Application US/09264604
Patent No. 6346411
GENERAL INFORMATION:

APPLICANT: HOSTETTER, MARGARET K.
APPLICANT: GALE, CHERYL A.
APPLICANT: BENDEL, CATHERINE M.
APPLICANT: TAO, NIAN-JUN
APPLICANT: KENDRICK, KATHLEEN
TITLE OF INVENTION: CANDIDA ALBICANS GENE, INTEGRIN-LIKE
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: MUEITING, RAASCH, GERHARDT & SCHWAPPACH, P. A.
STREET: 119 NORTH FOURTH STREET, SUITE 203
CITY: MINNEAPOLIS
STATE: MINNESOTA
COUNTRY: USA
ZIP: 55401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/264,604
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/642,846
FILING DATE: 03-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: MUEITING, ANN M.
REGISTRATION NUMBER: 33,977
REFERENCE/DOCKET NUMBER: 110.00280101
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-305-1217
TELEFAX: 612-305-1228
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1664 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-264-604-2

Query Match 5.0%; Score 109; DB 4; Length 1664;
Best Local Similarity 19.5%; Pred. No. 0.32;
Matches 80; Conservative 61; Mismatches 124; Indels 146; Gaps 20;

QY 92 NAAAGITRNOVALKAMARLARLARLAERDQPSISFPADLTSGRADPMVARAIADQOFTER 151
DB 974 NAKKGVTODEYI-----NAKLVDQPKKNSI-----VTDPEDRYEEL 1010

QY 152 ROT-----IOGVDMNNAORLOYOSEIEGIDRQGLKQDLGFIEDELIDLRK- 199
DB 1011 QOTASINMATIDSSYIGRPDISITDMLPYLSD-ELKKPTALLSADRLFMEOEVNRLRSN 1069

QY 200 -----LYDKGLVPRRLALLAERAGSLSGISGRLTADRKAAGVAGASDTOLKVRQI 249
DB 1070 SVLVHPGAGATNSMPEPDELINSPARVNSNSDNV-----AISGNAST-ISFNOL 1122

QY 250 KOEFFEQ--VSOSITE-----TVRLAEVTEKEVVASDAQ 282
DB 1123 DMNDDQATIQKIOEOPASKANTVRGDDGLASARETPRTPTKESISSKRAKLSAS 1182

QY 283 KR---IKIVSVNGTAQNLREFTEGAVVRAERLVDIAREDEAFVIAHFOPTVDN--V 337
DB 1183 PRKSPIKIGSV-----RVTKKNGSIAGIEPIPKATHKP-----KKSFGNGEISNKHV 1230

QY 338 HMGWTE-----VRLPAHSAGN---PDP-----ERND-- 362
DB 1231 RDGGISPSGSEHQHNPMSVSPSYQTDTATSTVDENKDVQYHPRKQOKHNNHNNHNN 1290

QY 363 -----PVAVADRISDP--OKQARLF--LGIIVRVQKOLRPHLGKRV 400

Db 1291 HHKORTDIPGVVDDEIPDVGLOEGRKLEFRRVLGKININLPDINTH-KGRFT 1340

RESULT 15

US-08-816-693A-51

; Sequence 51, Application US/08816693A
; Patent No. 5874241

GENERAL INFORMATION:

APPLICANT: Takahashi, Joseph S

APPLICANT: Turek, Fred W

APPLICANT: Pinto, Lawrence H

TITLE OF INVENTION: Clock Gene and Gene Product

NUMBER OF SEQUENCES: 53

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dressler, Rocky, Milnamow & Katz

STREET: Two Prudential Plaza, Suite 4700

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/816,693A

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: No. 5874241thrup, Thomas E

REGISTRATION NUMBER: 33,268

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-616-5400

TELEFAX: 312-616-5460

INFORMATION FOR SEQ ID NO: 51:

SEQUENCE CHARACTERISTICS:

LENGTH: 747 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-816-693A-51

Query Match 4.9%; Score 107; DB 2; Length 747;
Best Local Similarity 19.6%; Pred. No. 0.15;

Matches 72; Conservative 55; Mismatches 101; Indels 140; Gaps 17;

QY 86 LDPTQANA---AGIT-----RNQYVALK---AMEARLLAERDQPSISFP--- 125
DB 378 LEPRQFNLDSGLSPSSASSRSSHKSHTAMSEPISTPTKLMESTALPRATILPOL 437
QY 126 -----ADLTSGRADPMAVAIADE--OAOFTERRQTTGGVDIMNAQRIQ- 168
DB 438 PVGLSOAATMPLSSCDITQQLQPTLQSPAPQPSAQS--MFTIKDLE---QRTRI 492
QY 169 YQSEIEGIDROGKLDQGFIEDELIDRKLYDKGIVRPRRLALEARAGSLGSGIRL 228
DB 493 LDANIRMOOEELHKTIOEQLCLVQDSNVQ-----MFLQDPVSLSES----- 533
QY 229 TADRSKAVOGASDTOLKVKROIKEFFE--QVSOSITETRVRLAEVTE---KEVVASDAQ 282
DB 534 -----STORPAQOOLQORAAAPOLVLOGISTOYTOHLARESSVISGPKPMRSSOLS 587
QY 283 KRIKIVSPVNGTAQNLRFETBGAVVRAAEPLVDI---APEDEAFVIAHFOPTDVQNVHM 339
DB 588 GRSSSLSPEFSST-----LPPLITTPASTPOD----- 613
QY 340 GAVTEVRLPAFHSAGNPDERHDPVAVADRIDSPQKQARLFLGIVRVYDKOLPHLGRV 399
DB 614 -----SOCOPSPDFHD-----RQLRLLS-----QTIQPMKPGSC 643
QY 400 TAGMPAOV 407

Db 644 DAROPSEV 651

Search completed: April 22, 2003, 12:33:14
Job time : 23 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:27:47 ; Search time 35 Seconds

(without alignments)
2560.871 Million cell updates/sec

Title: US-09-913-414-5

Perfect score: 2165

Sequence: 1 MKPKRQRPDNDQAVARIG.....VLQYLFSPRLDRLTTRREE 435

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_todent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query length	DB ID	Description
1	1984.5	91.7	436	16	085351 caulobacter
2	1972.5	91.1	436	16	09rmn9 caulobacter
3	606	26.0	436	16	092nm2 rhizobium m
4	589	27.2	436	2	09x6n7 rhizobium l
5	589	27.2	474	16	08uaj1 agrobacteri
6	583	26.9	439	16	033679 rhizobium m
7	568	26.2	436	16	08u7n5 agrobacteri
8	564	26.1	443	16	09hyt9 pseudomonas
9	560.5	25.9	434	16	098l69 rhizobium l
10	543.5	25.1	437	16	p96438 rhizobium m
11	542.5	25.0	443	2	054417 serratia ma
12	536.5	24.8	443	2	005198 rhizobium m
13	530.5	24.5	511	16	092lt2 rickettsia
14	522	24.1	443	2	09xb63 erwinia amy
15	521.5	24.1	443	2	054457 serratia ma

17	519.5	24.0	511	16	092d4 rickettsia
18	512	23.6	435	2	005694 rhizobium l
19	510	23.6	433	2	09x693 pseudomonas
20	509	23.5	437	2	067994 pseudomonas
21	508.5	23.5	443	2	08vlt7 yersinia ru
22	502.5	23.2	447	16	0988e5 rhizobium l
23	499	23.0	435	2	09fcm8 rhizobium l
24	495	22.9	444	2	09znt8 pseudomonas
25	493	22.8	437	2	087809 pseudomonas
26	489.5	22.6	427	2	09zlg7 campylobact
27	482.5	22.3	427	2	09r8d9 campylobact
28	479.5	22.1	443	2	09rht1 pseudomonas
29	479	22.1	439	2	09r9h4 pseudomonas
30	477	22.0	438	2	09kg55 pseudomonas
31	476.5	22.0	442	16	08za92 yersinia pe
32	476.5	22.0	450	16	092ls8 rhizobium m
33	471.5	21.8	437	2	057387 serratia ma
34	467.5	21.6	440	16	098140 rhizobium l
35	467.5	21.6	441	2	085376 proteus mir
36	438.5	20.3	497	16	08uao0 agrobacteri
37	436	20.1	433	2	052859 rhizobium l
38	393	18.2	358	16	08x0n9 ralsontia s
39	375	17.3	387	16	08zmv7 salmonella
40	368.5	17.0	387	16	0824h8 salmonella
41	368	17.0	481	16	09kkm3 vibrio chol
42	367	17.0	395	16	0912m0 pseudomonas
43	329	15.2	475	16	09jy53 neisseria m
44	326.5	15.1	475	2	09x580 neisseria m
45	326.5	15.1	475	2	09jpk4 neisseria m

ALIGNMENTS

RESULT 1
ID 085351 PRELIMINARY; PRT; 436 AA.
AC 085351;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Membrane forming UNIT (R5AA secretion system, membrane protein R5AE).
GN R5AE OR CCI1009.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBL_TaxID=155892;
OX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15, AND NA1000;
RX MEDLINE=98292737; PubMed=9620954;
RA Avram P., Smit J.;
RT "The Caulobacter crescentus paracrystalline S-layer protein is secreted by an ABC transporter (Type I) secretion apparatus.";
RT J. Bacteriol. 180:3062-3069(1998).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15, AND NA1000;
RA Avram P., Smit J.K.;
RT Submitted (Apr-1998) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15, AND NA1000;
RA Avram P.;
RT Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.
[4]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laud M.T., Paulsen I.T., Nelson K.E., Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R., Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,

RA Deboy R.T., Dodson R.J., Durkin A.S., Gwin M.L., Haft D.H.,
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
 RA Ueberlack T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.,
 RT "Complete genome sequence of *Caulobacter crescentus*,"
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 DR EMBL: AF062345; AAC38667.2;
 DR EMBL: AF062345; AAC38667.2;
 DR TIGR: CC1009;
 DR InterPro: IPR002215; HLYD.
 DR InterPro: IPR003997; REXD.
 DR Pfam: PF00529; HLYD.1
 DR PRINTS: PR01490; RYTXOXIND.
 DR PROSITE: PS00543; HLYD_FAMILY; UNKNOWN.1.
 DR Complete proteome.
 SQ SEQUENCE 436 AA; 48389 MM; 90E32A834673C9DC CRC64;

Query Match 91.7%; Score 1984.5; DB 16; Length 436;
 Best Local Similarity 93.6%; Pred. No. 1,3e-116;
 Matches 408; Conservative 3; Mismatches 24; Indels 1; Gaps 1;

QY 1 MKPRTIQPTDFQAVARIGYIIALTFVGLGMAAFAPLDSAVIANGVSAEVS-ODVQ 59
 DB 1 MKPRTIQPTDFQAVARIGYIIALTFVGLGMAAFAPLDSAVIANGVSAEGRKTYQ 60
 QY 60 HLEGMLAKILVREGEKVKAGOVLELDPQTQANAAAGITRNQYALKAMEARLLAERDQ 119
 DB 61 HLEGMLAKILVREGEKVKAGOVLELDPQTQANAAAGITRNQYALKAMEARLLAERDQ 120
 QY 120 PSISFPADLTSGRADPMVARRAIADQAOFTERRQTIQGVDMNORLOYOSEIEGIDRQ 179
 DB 121 PSISFPADLTSGRADPMVARRAIADQAOFTERRQTIQGVDMNORLOYOSEIEGIDRQ 180
 QY 180 TQGLKDQGFTEDELIDRLKLYDKGLVPRRLALEARAGSLSGISGRLLADRASKAVOGA 239
 DB 181 TQGLKDQGFTEDELIDRLKLYDKGLVPRRLALEARAGSLSGISGRLLADRASKAVOGA 240
 QY 240 SDTQKVKQIKOEFEQVQSSTTERVRLAEVTEKEVVASDQAKRIKIVSPVNGTAQMLR 299
 DB 241 SDTQKVKQIKOEFEQVQSSTTERVRLAEVTEKEVVASDQAKRIKIVSPVNGTAQMLR 300
 QY 300 FETEGAVVRAAEPLVDIAPEDAEFYIOAHFOPTVDVNMGMVTEVRLLPAFHSAGNPDP 359
 DB 301 FETEGAVVRAAEPLVDIAPEDAEFYIOAHFOPTVDVNMGMVTEVRLLPAFHSAGNPDP 360
 QY 360 RHDPAVADRISDPQKARLEGLIVRVYKQLPRLRGRTVAGMPAOYIVPTGERTVLYQ 419
 DB 361 GTIOSLSODRISDPQKARLEGLIVRVYKQLPRLRGRTVAGMPAOYIVPTGERTVLYQ 420
 QY 420 LFSPLRDLRTTMRRE 435
 DB 421 LFSPLRDLRTTMRRE 436

RESULT 2

Q9RMM9 PRELIMINARY; PRT; 436 AA.
 AC Q9RMM9;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Membrane forming unit.
 GN R5AF.
 OS *Caulobacter crescentus*.
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
 OC Caulobacter.
 OX NCBI_TaxID=155892;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=J54050;
 RT Bingle W.H., Avram P.A., Nornellini J.F., Smit J.K.;
 RT "The secretion signal of the Caulobacter crescentus S-layer protein is
 located within the C-terminal 82 amino acids of the molecule."

RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF193064; AAF07963.1;
 DR InterPro: IPR002215; HLYD.
 DR InterPro: IPR003997; REXD.
 DR Pfam: PF00529; HLYD.1
 DR PRINTS: PR01490; RYTXOXIND.
 DR PROSITE: PS00543; HLYD_FAMILY; UNKNOWN.1.
 DR SEQUENCE 436 AA; 48443 MM; F9C4C680F3D96563 CRC64;

Query Match 91.1%; Score 1972.5; DB 2; Length 436;
 Best Local Similarity 93.1%; Pred. No. 7.4e-116;
 Matches 406; Conservative 3; Mismatches 26; Indels 1; Gaps 1;

QY 1 MKPRTIQPTDFQAVARIGYIIALTFVGLGMAAFAPLDSAVIANGVSAEVS-ODVQ 59
 DB 1 MKPRTIQPTDFQAVARIGYIIALTFVGLGMAAFAPLDSAVIANGVSAEGRKTYQ 60
 QY 60 HLEGMLAKILVREGEKVKAGOVLELDPQTQANAAAGITRNQYALKAMEARLLAERDQ 119
 DB 61 HLEGMLAKILVREGEKVKAGOVLELDPQTQANAAAGITRNQYALKAMEARLLAERDQ 120
 QY 120 PSISFPADLTSGRADPMVARRAIADQAOFTERRQTIQGVDMNORLOYOSEIEGIDRQ 179
 DB 121 PSISFPADLTSGRADPMVARRAIADQAOFTERRQTIQGVDMNORLOYOSEIEGIDRQ 180
 QY 180 TQGLKDQGFTEDELIDRLKLYDKGLVPRRLALEARAGSLSGISGRLLADRASKAVOGA 239
 DB 181 TQGLKDQGFTEDELIDRLKLYDKGLVPRRLALEARAGSLSGISGRLLADRASKAVOGA 240
 QY 240 SDTQKVKQIKOEFEQVQSSTTERVRLAEVTEKEVVASDQAKRIKIVSPVNGTAQMLR 299
 DB 241 SDTQKVKQIKOEFEQVQSSTTERVRLAEVTEKEVVASDQAKRIKIVSPVNGTAQMLR 300
 QY 300 FETEGAVVRAAEPLVDIAPEDAEFYIOAHFOPTVDVNMGMVTEVRLLPAFHSAGNPDP 359
 DB 301 FETEGAVVRAAEPLVDIAPEDAEFYIOAHFOPTVDVNMGMVTEVRLLPAFHSAGNPDP 360
 QY 360 RHDPAVADRISDPQKARLEGLIVRVYKQLPRLRGRTVAGMPAOYIVPTGERTVLYQ 419
 DB 361 GTIOSLSODRISDPQKARLEGLIVRVYKQLPRLRGRTVAGMPAOYIVPTGERTVLYQ 420
 QY 420 LFSPLRDLRTTMRRE 435
 DB 421 LFSPLRDLRTTMRRE 436

RESULT 3

Q92NX2 PRELIMINARY; PRT; 436 AA.
 AC Q92NX2;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Putative toxin secretion transmembrane protein.
 GN R02036 OR SMC04208.
 OS *Rhizobium melioidi* (*Sinorhizobium melioidi*).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; *Sinorhizobium*.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RC MEDLINE=21396507; PubMed-11481430;
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
 RA Boistard D., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masny D.,
 RA Pohl T., Portet-Berle D., Pehler A., Punnelle B., Ramsperger U.,
 RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.;
 RT "Analysis of the chromosome sequence of the legume symbiont
Sinorhizobium melioidi strain 1021."
 RT Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
 DR EMBL: AL591789; CAC46615.1;
 DR InterPro: IPR002215; HLYD.

ID	ORGANISM	PRELIMINARY	PR	AA
AC	09X6N7			
AD	01-NOV-1999 (TREMblrel. 12, Created)			
DT	01-NOV-1999 (TREMblrel. 12, Last sequence update)			
DT	01-DEC-2001 (TREMblrel. 19, Last annotation update)			
DE	Rhizobium secretion protein RspE.			
GN	RspE.			
OS	Rhizobium leguminosarum (biovar trifolii).			
OC	Plasmid pRhe162Y10C.			
OC	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;			
OC	Rhizobiaceae; Rhizobium.			
OX	NCBI_TaxID=386;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=162Y10;			
RC	MEDLINE=99318631; PubMed=10388672;			
RT	Oresnik I.J., Twelker S., Hynes M.F.;			
RT	"Cloning and characterization of a Rhizobium leguminosarum gene			
RT	encoding a bacteriocin with similarities to RTX toxins.,"			
RT	Appl. Environ. Microbiol. 65:2833-2840(1999).			
RL	[12]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=162Y10;			
RC	MEDLINE=21360352; PubMed=11467725;			
RT	Venter A.P., Twelker S., Oresnik I.J., Hynes M.F.;			
RT	"Analysis of the genetic region encoding a novel rhizobiocin from			
RT	Rhizobium leguminosarum by virAlike strain 306.,"			
RL	Can. J. Microbiol. 47:495-502(2001).			
DR	EMBL; AF141933; AAD4891.1; -			
DR	InterPro; IPR002215; HLyd.			

RESULT 5			
080UAJ1			
ID	080UAJ1	PRELIMINARY;	PRT; 474 AA.
AC	080UAJ1;		
DT	01-JUN-2002 (TrEMBLrel. 21, Created)		
DT	01-JUN-2002 (TrEMBLrel. 21, last sequence update)		
DT	01-JUN-2002 (TrEMBLrel. 21, last annotation update)		
DE	Secretion protein, HlyD family.		
GN	PRE OR ATU3382 OR AGR_L2881.		
OS	Agrobacterium tumefaciens (strain C58 / ATCC 33970).		
OC	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;		
OC	Rhizobiaceae; Rhizobium.		
OX	NCBI_TaxID=176289;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	Wood D.W., Setubal J.C., Kau L., Monks D.E., Kitajima J.P.,		
	Wood D.W., Setubal J.C., Kau L., Monks D.E., Kitajima J.P.,		
	Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,		
	Chen Y., Paulsen I.T., Eisen J.A., Kap P.D., Boeve D. Sr.,		
	Chapman P., Clendenning J., Deatherage G., Gilllet W., Grant C.,		
	Ra Kityavain T., Levy R., Li M.-J., McClelland E., Palmeri A.,		
	Ra Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,		
	Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan M., Perry M.,		
	Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,		
	Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,		
	Nester E.W.;		
RT	"The genome of the natural genetic engineer Agrobacterium tumefaciens		
RL	C58."		
RN	Science 294:2317-2323(2001).		
	[2]		

[illegible]

RL Mol. Microbiol. 25:117-134(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-1021; PLASMID-PSYMB (MEGAPLASMID 2);
 RX MEDLINE=21396508; PubMed=11481431;
 RA Fijan T.M., Weidner S., Wong K., Buhmester J., Chain P.,
 RA Vorhoefer F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
 RA Golding B., Puelher A.;
 RT "The complete sequence of the 1,683-bp *psymb* megaplasmid from the N2-
 RT fixing endosymbiont *Sinorhizobium meliloti*."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
 DR EMBL; U89163; AAB64092.1;
 DR EMBL; AL603646; CAC49691.1;
 DR InterPro: IPR002215; HLYD.
 DR InterPro: IPR003997; REXD.
 DR Pfam: PF00529; HLYD.1
 DR PRINTS; PR01490; RTYTOXIND.
 KW Plasmid; Complete proteome.
 SO SEQUENCE 439 AA; 48036 MW; 82B2421825522BD CRC64;
 Query Match 26.9%; Score 583; DB 16; Length 439;
 Best Local Similarity 31.9%; Pred. No. 9,2e-29;
 Matches 134; Conservative 84; Mismatches 199; Indels 4; Gaps 4
 QY 19 IGYIATLTV-GLIGMAAPLDSAVTANGVSAEVS-QDVQHLBEGMLAKILYREGK 76
 DB 21 IGVSVLATALVAGVGGMATTELSAIVAGVVIYDDNKKYQHLGLGIVGELLVKEGDR 80
 QY 77 VKAGOVLEFLDPTQNAAGGIRTRNOYVALKANEARLARODRPSISFPADLTSGRADPM 136
 DB 81 VKAGOVLLRLDSTTVKRAVNAITESTLAQYKARRAKQAEKMGAASTFIEDLAETLPGR 140
 QY 137 VARAIADQAOFTERRQTIQGVQVNDLMAQRLQYQSEIEGIDRQGLKQDLGTEDELID 196
 DB 141 AAKLIEGEQRLTASRRSALSGKMGKQDLSKQALADVEVEELTVQLNAIEEALKLIAEELTG 200
 QY 197 LKLYDKGIVPRPRLLEARAGSGSGSLGLTRADRSKAVQGSADPQLKVRDQKQEFFQ 256
 DB 201 VDSLGGQGLVPRMQRVTTLLKQRAELLEGGRHIAAQAQKSSSELDLQDQDEDRSE 260
 QY 257 VSQSTETFEVRRLAEVTEKEEVASDAQKRLKIVSPVNGTAQNLFFTEGAVVRAAEPLVDI 316
 DB 264 ISKELTLDVEAKIAEYEBERTATDQLRRDLITAPLSGRITYOLAIRHVNGVINPGETLMV 320
 QY 317 APEDEAFVYQAFQPTDVUNVHMGAVTEVRLAFHSAQNPDEBRDPVAVADRISPPQK 376
 DB 321 VEAEADLVYEAQVATHTDIDQIRGVGSVEIRFSAFMORTPEVEAEVYVAPLVYDERG 380
 QY 377 ARFLFGIVVDKOLPPLHKG-RVTAQMPAQVIVPGERTVQVLESPRLTLRTMRRE 435
 DB 381 ASYVPLRIRPKASL-AKLKGLSLVPGMAEAFYIKIADTVISYLTKPLTDQMRNAFRD 439
 RESULT 7
 ID 0807N5 PRELIMINARY; PRT; 436 AA.
 AC 0807N5;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE ABC transpositer, membrane spanning protein.
 GN R39E OR ATU4414 OR AGR_L_906.
 OS *Agrobacterium tumefaciens* (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizodium.
 OX NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608550; PubMed=11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eissen J.A., Karp P.D., Boyce D., Sr.,
 RA Chapman P., Glendening J., Deatherage G., Gillet W., Grant C.,

Pseudomonas aeruginosa.

Phyllobacteriaceae; Mesorhizobium.

OX NCBI_TaxID=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-MAFF303099;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneo T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida T., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 Mesorhizobium loti";
 RL DNA Res. 7:331-338(2000).
 DR EMBL, AF002996; BAB48494.1;
 DR InterPro: IPR002106; ATRNA_LigaseII.
 DR InterPro: IPR002215; HlyD.
 DR InterPro: IPR003997; RtxD.
 DR Pfam: PF00529; HlyD; 1.
 DR PRINTS: PR01490; RXTXOIND.
 DR PROSITE, PS00339; AA_TRNA_LIGASE_IL_2; UNKNOWN_1.
 DR Complete proteome.
 KW SEQUENCE 434 AA; 46926 MW; 07BD024352BD919 CRC64;
 SO
 Query Match 25.9%; Score 560.5; DB 16; Length 434;
 Best Local Similarity 35.1%; Pred. No. 2.3e-27;
 Matches 148; Conservative 68; Mismatches 193; Indels 13; Gaps 6;
 OY 22 GIATLETV--GLGMAFAPLDSAVIANG--VSAEVSQDYOHLEGMARILVREGKV 77
 DB 18 GVAACIFLVGAGSLAAVTELSGAVIAPKLVDSYVK--VQHPGCVGVDILAREGDAV 76
 OY 78 KAGVLEFLDPTQANNAAGITRNOYVALKAMEARLARERDQPSISFPADITSORADMY 137
 DB 77 KSGOVLRLDELTVTRANLAIVTKGLDEFEARLARLEARDRAGIAFPASLTSRDPAV 136
 OY 138 ARAIADDAQETTEROTIGOVDMNAORLYOSEIEGIDRQGLKQGLFIEDELIDL 197
 DB 137 ARAAAGEOSLEFERNQAGAKQOLEERIALAEASGLTQRKAKSEIELIGLELSI 196
 OY 198 RKLIDGLVPRRLALEARAGSLSGISGRLTADRSKAVQASDTOLKVRQIKOEFBOV 257
 DB 197 RTLMKRLIVSIDRMALERDAVRLDGEHQGLTASIAQSKGRILAETRLDIIQVDDLRSEV 256
 OY 258 SOSITERVRRLAEVTEKEVNASDAOKRIKITYSPVNGTQONLRFLEGANVRAAEPLVIA 317
 DB 257 ATELADVGKISLEVERKVSADOLKTRIDRSPQGVVHQAIVHTIGVISPEGVIMLV 316
 OY 318 PEDBAFVIOAHFOPTDVNVHMGWTEVRLPAFHSAGNPDEPRHDPAVADRISDPKQA 377
 DB 317 PVADDLVEARIALAFODIQLSLGQDVALKLSAFNQRVTPPELSGVSEISADLSVDERGA 376
 OY 378 RLF---LGIIVNDVKQLPRLRGVTA--GMPAOYIVPTGERTVLOYLEFSPRLDRLTMR 433
 DB 377 SFYVVRVSLPRLTELEK----LKGLTLAPGMVEAFATGSTRMLSYLKLADQARAFR 432
 OY 434 EE 435
 DB 433 EE 434

RESULT 10
 P96438 PRELIMINARY; PRT; 473 AA.
 AC P96438;
 DT 01-MAY-1997 (Tremblrel. 03, Created)
 DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE EXPD2 (Hypothetical protein).
 GN EXPD2 OR R80891 OR S8B21315.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Plasmid psymb (megaplasmid 2).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.

OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-RCR2011 / SU47;
 RX MEDLINE=91175570; PubMed=9023225;
 RA Becker A., Rueberg S., Kuester H., Roxlau A.A., Keller M.,
 RA Iwashina T., Cheng H., Walker G.C., Puhler A.;
 RT "The 32-kilobase exp gene cluster of Rhizobium meliloti directing the
 RT biosynthesis of galactoligucan: genetic organization and properties of
 RT the encoded gene products";
 RL J. Bacteriol. 179:1375-1384(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-1021; PLASMID-PSYMB (MEGAPLASMID 2);
 RX MEDLINE=21396508; PubMed=11481431;
 RA Flann T.M., Weidner S., Wong K., Buhrmaster J., Chain P.,
 RA Vorholter F.J., Hernandez-Lucas I., Becker A., Cowle A., Gouy J.,
 RA "The complete sequence of the 1,683-kb psymb megaplasmid from the N2-
 RT fixing endosymbiont Sinorhizobium meliloti";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
 DR EMBL; Z79692; CAB01939.1;
 DR EMBL; AL603645; CAC49291.1;
 DR InterPro: IPR002215; HlyD.
 DR InterPro: IPR003997; RtxD.
 DR Pfam: PF00529; HlyD; 1.
 DR PRINTS: PR01490; RXTXOIND.
 DR Hypothetical protein; Plasmid; Complete proteome.
 KW SEQUENCE 473 AA; 51479 MW; C0C3B168D9C23D85 CRC64;
 SO
 Query Match 25.1%; Score 543.5; DB 16; Length 473;
 Best Local Similarity 28.8%; Pred. No. 3e-26;
 Matches 126; Conservative 97; Mismatches 207; Indels 7; Gaps 3;
 OY 4 PKIOPTEPQAVAR---IGYIATLFVGLGMAFAPLDSAVIANGVSAEVS--QDV 58
 DB 39 PEPLSEBNTSPRLRLVIANLITLIVAFGFGFMATSTELSSASVSGTIVDSKRTV 98
 OY 59 QHLEGMLAKTLVBEGEKVAGQVLEFDPQANNAAGITRNOYVALKAMEARLARERDQ 118
 DB 99 SHFEGVIGRLVQEGDHVAPGQPLMKLEDTRASDLOALQSRVGLAKLARAEARAG 158
 OY 119 RPSISFPADLTSORADPVARAIADEQAOTFERRQTIQGOVDLMNAORLYQOSEIEGDR 178
 DB 159 LHAVSFPADLSE--GEAAADAVTAERKAFFERSEAESRLAIQRTIEEYSEKASLTA 216
 OY 179 QTOGLKQGLFIEDELIDRLKLYDGLVPRRLALEARAGSLSGISGRLTADRSKAVOG 238
 DB 217 QLOATDRQIELMNOQRRIATLVEKAFQRSKLIEIDARLELATGELAGDAQAQAKA 276
 OY 239 ASDTOLKVRQIKOEFEEVSOSITETRVRLAEVTEKEVNASDAOKRIKITYSPVNGTQONL 298
 DB 277 MAGAELALIGIESDFOSRIAGEITFARLEAEVQRITAADVLRRLREIRAPQAGIVANI 336
 OY 299 RFTLEGANVRAAEPLVNDAPDEAFVIOAHFQPPDVNVHMGWTEVRLPAFHSAGNPDP 358
 DB 337 QLRTPGSVAIVPGPLDIPDEPLVEMHVSITRDIDISITIGSTQRLTAVNRSHPPL 396
 OY 359 ERHDPVAVADRISDPKQARLELGIIVNDVKQLPRLRGVTA--GMPAOYIVPTGERTVLO 418
 DB 397 EGYKTYIAADQSVDEKSNVAIFVARAEVTPESLANANDIRILYPCMPAEVLIVHRSRAID 456
 OY 419 YLFSPLDRLTFTTMR 435
 DB 457 YLVAPVSDSFNARFRED 473

RESULT 11
 Q8UBZ2 PRELIMINARY; PRT; 437 AA.
 AC Q8UBZ2;
 DT 01-JUN-2002 (Tremblrel. 21, Created)
 DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE HlyD family secretion protein.
 GN PRSE OR ATU2706 OR AGR_C_4904.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608550; PubMed=11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boyce D. Sr.,
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
 RA Kutyavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
 RA Raymond C., Rouse G., Stenphimachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nester E.W.;
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
 C58.";
 RL Science 294:2317-2323(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Ouyollo B., Goldman B.S., Cao Y., Askenzi M., Halling C., Mullin L.,
 RA Houmel K., Gordon J., Vaudin M., Tatchouk O., Epp A., Liu F.,
 RA Wollam K., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
 RA Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
 RA Cielo C., Slater S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent
 Agrobacterium tumefaciens C58.";
 RL Science 294:2323-2328(2001).
 DR EMBL; AE009216; AAL3687.1; -
 DR EMBL; AE008183; AAK88425.1; -
 KW Complete proteome.
 SQ SEQUENCE 437 AA; 4773 MW; 67479F1D40122D04 CRC64;

Query Match 25.1%; Score 542.5; DB 16; Length 437;
 Best Local Similarity 30.8%; Pred. No. 3.2e-26;
 Matches 130; Conservative 83; Mismatches 204; Indels 5; Gaps 2;
 Oy 15 AAVAGYGIATLFGGLGMAFAPLDSAVINGVSAE-VSODVOHLEGMIAKILVREGK 73
 Db 20 AVVVLGLGLL-VGMGMAFAPLDSAVINGVSAE-VSODVOHLEGMIAKILVREGK 75
 Oy 74 GERVAGVLEFELPTQANAAGITRNOYVALKAMEARLLAERDORPSISPPADLTSGRA 133
 Db 76 GDRVAGVLEFELPTQANAAGITRNOYVALKAMEARLLAERDORPSISPPADLTSGRA 135
 Oy 134 DPMVARLADDEQAOTERRQITQGVODLMMNORLQYOSIEIDRQGLKQOLGFIED 193
 Db 136 SKSARFTIDSEONLNSRNALITGMKKOLATRRQLADEARGLDQVTEENELAIYKED 195
 Oy 194 LIDLRLKLDKGVPPRLALAEARAGSLGSGRLTADRSKAVQASDPQLVROIKOFE 253
 Db 196 VSKTEBLAKKGLVLTORLNLKRLQSLNEGOGQYIARAQYVGLSELDLQDLQDER 255
 Oy 254 FEQVSOSITETRVRLAEVTEKEVNASDAQRIKIVSPVNGTAONLRFETEGAVVRAEPL 313
 Db 256 KSEVTKDLTSIRATVAEEERLAATRDQLDRDISPIGRITQYSVNHINCVIDPEVL 315
 Oy 314 VDIADDEAFVIOAHFOPIDVDVNHMGVTEVRLPAFHSAGNPDEPRHPVAVADRISP 373
 Db 316 MLVVDKDLAEANITPPDDIIVGQPVIVRTAFNGSTPDLASAEEVAVAPDLQDTS 375
 Oy 374 OKQARLFLGIVAVDKQPLPHLGRVHTAGMPAIVPTGERTVLYLFESPLDRTTMR 433
 Db 376 RGTSTYVLRIRPNKAGMHLPGKILYFGMPAEVIGTISERSVLYFYKPFDDRLKFTV 435
 Oy 434 EE 435

Db 436 OE 437
 RESULT 12
 ID 054417 PRELIMINARY; PRT; 443 AA.
 AC 054417;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Metalloprotease transporter.
 GN HASB8000.
 OS Serratia marcescens.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Serratia.
 OX NCBI_TaxID=615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SR41;
 RA Omori K.;
 RT "Cloning of the hasB8000 genes from Serratia marcescens Str41:
 RT secretion mechanism for the lipase and metalloprotease.";
 RL Submitted (FEB-1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL; D83582; BAA12016.1; -
 DR InterPro: IPR002215; HLYD.
 DR InterPro: IPR003997; RtxD.
 DR Pfam: PF00529; HLYD; 1.
 DR PRINTS; PRO1450; RYTOXIND.
 DR PROSITE; PS00543; HLYD_FAMILY; 1.
 KW Protease.
 SQ SEQUENCE 443 AA; 48409 MW; 762FD7F822B0C44 CRC64;

Query Match 25.0%; Score 540.5; DB 2; Length 443;
 Best Local Similarity 31.7%; Pred. No. 4.3e-26;
 Matches 138; Conservative 85; Mismatches 178; Indels 35; Gaps 6;
 Oy 18 RIGYGIATLFGGLGMAFAPLDSAV-TANCVSAEVSODVOHLEGMIAKILVREGK 76
 Db 25 RIGGMVILVFGFGLMGLAPLDKGVPSGVSVAAGNKKAVQHPGSGVSOIQVHEGR 84
 Oy 77 VKAGVLEFELPTQANAAGITRNOYVALKAMEARLLAERDORPSISPPADLTSGRADP 136
 Db 85 VVAGVLEFELPTQANAAGITRNOYVALKAMEARLLAERDORPSISPPADLTSGRADP 144
 Oy 137 VARLADDEQAOTERRQITQGVODLMMNORLQYOSIEIDRQGLKQOLGFIED 196
 Db 145 VMSLMLOQLFTSRRAALQSELAIAESIAQSQAVLEGRISYASKOROKAMLOEOLG 204
 Oy 197 LRLKLDKGVPPRLALAEARAGSLGSGRLTADRSKAVQASDPQLVROIKOFE 249
 Db 205 MKRLAAGVVAANRLDLDEGOYAIDGQASDPTGNLRL-----GRQLLEKLRAI 255
 Oy 250 --KQEFFEVSOSITETRVRLAEVTEKEVNASDAQRIKIVSPVNGTAONLRFETEGAVV 307
 Db 256 QRREYQKESVQSLEVRKLDLDELNRLAKAEDLGHQYKASVAGTVVGLVFTGEGVI 315
 Oy 308 RAEPVLDLADDEAFVIOAHFOPIDVDVNHMGVTEVRLPAFHSAGNPDEPRHPVANA 367
 Db 316 GAGQQLMETVPSDRGLQVBARIPVELIDKVOGLVPELLEFSAFNPSTPRVGEVTLVA 375
 Oy 368 DRISDPKQARL-----LGIIVAVDKQPLPHLGRVHTAGMPAIVPTGERTVLY 419
 Db 376 DRLTEKSGAPYYSRAKVSAGLQRLNLEIRP-----GMPEGRITGSGSMNY 427
 Oy 420 LFSPLDRLTTRREE 435
 Db 428 LFKPLDRLHLLATTE 443
 RESULT 13
 ID 005198 PRELIMINARY; PRT; 473 AA.

AC 005198: (Tremblrel. 04, Created)
 DT 01-JUN-1997 (Tremblrel. 12, Last sequence update)
 DT 01-NOV-1999 (Tremblrel. 17, Last annotation update)
 DE Membrane fusion protein.
 GN EXPD2.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ERR1:
 RA MEDLINE=98162063; PubMed=9501442;
 RA Lioret J., Wolff B.B.H., Rubio J.M., Downie J.A., Bonilla I.,
 RA "Exopolysaccharide II production is regulated by salt in the
 RT halotolerant strain Rhizobium meliloti EFB1."
 RL Appl. Environ. Microbiol. 64:1024-1028(1998).
 DR EMBL: Y08703; CAB41456.1; -
 DR InterPro: IPR002215; HLYD.
 DR InterPro: IPR003997; RLXD.
 DR Pfam: PF00529; HLYD: 1.
 DR PRINTS: PR01490; FTXTOXIND.
 SO SEQUENCE 473 AA; 51603 MW; D536469E385319FF CRC64;

Query Match 24.8%; Score 536.5; DB 2; Length 473;
 Best Local Similarity 28.5%; Pred. No. 8.4e-26;
 Matches 119; Conservative 96; Mismatches 199; Indels 3; Gaps 2;

QY 20 GYGIALLTFVLLGMAFAPLDSAVIANGVSAEVS-QDVHLEGGMALIKLVGEVEYK 78
 DB 59 GLTTLVAFGGFGMAFSTELSSASVSTGTLIVDSKRTVSHFEGVLSRLIVOGDHVA 118
 QY 79 AGVLFELDPQANAAAGITRNOYVALKAMEARLAEQPSISFADLTQSRADPVVA 138
 DB 119 PGGLMLQEDTRASDLOALSRRVGLAKLARLSLAGIQAADFDDVA--AGEAAA 176
 QY 139 RAIDEGAOFTERRQITOGVDLMAAORLOYSEIEGIDRQTGICKDGLFIEDELIDLR 198
 DB 177 DAVAEAFEFKRNBAKRGRIQRTIEEYSEKAKSLTAQLQATDRQIELMNEQRATA 236
 QY 199 KLYRKGIVPRRLALERAGSLSGSLTADSKAVOGASDPLQKROTKOEFEEVYS 258
 DB 237 TLVKAFAQRSKLEIRARSELATRGELAGRAQAKMAAGALTLTGIESPFOSEA 296
 QY 259 QSIETFRVRLAEVTEKEVYASDAQRIKIVPVNGTAONLFEFEGAVVRAEPLVDIAP 318
 DB 297 GEITTALELAEEVERIISAKDVLRLERAPQAGIVANIQLRTPGSAVTGQPLDILV 356
 QY 319 EDEAFVIOAHFOPTVDNVHMGVTEVRLPAFHSAGNDPERRHDPVAVADRISPOKAR 378
 DB 357 EDEPLLVEMHVSITRDIDISITIGSSQIRLTAYNQRSHLPMGKVTYIAADQSMDEKNVA 416
 QY 379 LELGIVAVDVQKLPRLGRVTAGMPAVIYPTGERVLYQLFSLRDLTMTAREE 435
 DB 417 YFVARAEIAPESLANPDIRLYPGMPAEVLIVHAKARSAIDYLVSPDSFNARRED 473

RESULT 14
 0921J2
 ID 0921J2 PRELIMINARY; PRT; 511 AA.
 AC 0921J2;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Alkaline protease secretion protein AprE.
 GN APRE OR KC0428.
 OS Rickettsia conorii.
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.
 OX NCBI_TaxID=781;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=MALISH 7;
 RX MEDLINE=21442074; PubMed=11557893;
 RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
 RA Samsom D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
 RA Roult D.;
 RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii";
 RL Science 293:2093-2098(2001).
 DR EMBL: AE008606; AAL02966.1; -
 DR InterPro: IPR002215; HLYD.
 DR Pfam: PF00529; HLYD: 1.
 DR PROSITE: PS00543; HLYD_FAMILY; UNKNOWN_1.
 KW protease; Complete proteome.
 SO SEQUENCE 511 AA; 57582 MW; 4BD98A55866A5C CRC64;

Query Match 24.5%; Score 530.5; DB 16; Length 511;
 Best Local Similarity 30.0%; Pred. No. 2.2e-25;
 Matches 135; Conservative 88; Mismatches 202; Indels 25; Gaps 7;

QY 5 KIQPTDNEQAVAR-IGYGIALLTFVGLG--MAAFAPLDSAVIANGVSAEVS-QDVQ 59
 DB 68 KTDKDRNNVAQAARSPILFGIYVIFVLIGLSALAPLDSGAVAVGIVMPSTNKTIQ 127
 QY 60 HLEGGMALIKLVREGGEYKAGQVLELDPQANAAAGITRNOYVALKAMEARLAEQDQR 119
 DB 128 HHEGGIINAIYVKGQDKKEDKLELETRIKSEHENILGQYRNFATENRLAERDNL 187
 QY 120 PSIFPADLTQSRADPVVRAIADQAOFTERRQITOGVDLMAAORLOYSEIEGIDRQ 179
 DB 188 EQIESFSPDLMDINLPEVAKITHTQENLFRSRKEYNSEKALQNTAQLEKKEIGLEAK 247
 QY 180 TQGLKDGLFIEDELIDLRKLYDGLVPRRLALERAGSLSGSLTADSKAVOGA 239
 DB 248 KVAASKTAEEYQDRKLALRTLEKEFGVQKALLDQEAQVAAKSDVATTEAIGIRHAI 307
 QY 240 SDTQKVRQIQEPEEVOVSQSIETFRVRLAEVTEKEVYASDAQRIKIVPVNGTAONLR 299
 DB 308 TETQIKIINQNKYTERLTLELREAOVQFASLKERYNALDLSLRVITIRAVDAIVSLK 367
 QY 300 FFEQAVVRAEPLVDIAPDEDAFVIOAHFOPTVDNVHMGVTEVRLPAFHSAGN--- 356
 DB 368 YHTIGCVISHQOPTMEISPTNDPLIEKVSQKNDSHESLVAKIRSAKSTTPTPT 427
 QY 357 -DPERHDPVAVADRISDP-QKQARLFLGIVRD-----VKQPLRLGRVTAGMPA 405
 DB 428 GKVVSISSPDIVQDERQPGQGDQNVYAVARVIEDDFENKVAKVKNLEH-----PGMQA 481
 QY 406 QVIYPTGERVLYQLFSLRDLTMTAREE 435
 DB 482 EVQIVTGTIRLLRYLDELVDVTDAERAREK 511

RESULT 15
 09XB63
 ID 09XB63 PRELIMINARY; PRT; 443 AA.
 AC 09XB63;
 DT 01-NOV-1999 (Tremblrel. 12, Created)
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE PrtE protein.
 GN PRTE.
 OS Erwinia amylovora.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Erwinia.
 OX NCBI_TaxID=552;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=E9;
 RA MEDLINE=99303693; PubMed=10373365;
 RA Zhang Y., Bak D.D., Heid H., Geider K.;
 RT "Molecular characterization of a protease secreted by Erwinia
 RT amylovora";
 RL J. Mol. Biol. 289:1239-1251(1999).

THIS PAGE BLANK (per 10)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:25:06 ; Search time 25 Seconds

(without alignments)
721.688 Million cell updates/sec

Title: US-09-913-414-5

Perfect score: 2165

Sequence: 1 MKPPKIQRPDNGAVARIG.....VLQYLFSPRLRLTKTRKEE 435

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	514.5	23.8	432	1	APRE_PSEAE
2	506.5	23.4	448	1	PTE_ERWCH
3	313.5	14.5	477	1	RT3D_ACTPL
4	288.5	13.3	478	1	HLV4_ECOLI
5	267.5	12.4	477	1	HLVD_ACTAC
6	267.5	12.4	478	1	HLVD_ECOLI
7	266	12.3	440	1	CYAD_BORPE
8	243.5	11.2	478	1	HLVD_PASBP
9	233.5	10.8	477	1	RT1D_ACTPL
10	233	10.8	478	1	MCHE_PASHA
11	164	7.6	424	1	MCHE_ECOLI
12	163	7.5	413	1	CYAA_ECOLI
13	147.5	6.8	343	1	YTCR_ECOLI
14	146	6.7	421	1	EXOF_RHIME
15	142	6.6	355	1	YH1I_ECOLI
16	136.5	6.3	382	1	YB94_HAEIN
17	130.5	6.0	879	1	MSP_ONCVO
18	130	6.0	1625	1	CTPI_MYCTU
19	129.5	6.0	848	1	MSP_DIRM
20	128	5.9	371	1	MACA_ECOLI
21	128	5.9	882	1	MSP_CAEEL
22	126.5	5.8	331	1	YBHG_SALTY
23	126	5.8	390	1	EMRA_ECOLI
24	125	5.8	331	1	YBHG_SALTY
25	124	5.7	371	1	MACA_YERPE
26	124	5.7	880	1	MSP_BRUMA
27	123.5	5.7	1576	1	POC_AQUY
28	122.5	5.6	378	1	YIBH_ECOLI
29	122	5.6	334	1	YH3_RHIME
30	122	5.6	371	1	MACA_ECOS7
31	121.5	5.6	378	1	YIAY_ECOLI
32	121.5	5.6	545	1	HTR5_HALNI
33	121.5	5.6	4684	1	PLE1_HUMAN

ALIGNMENTS

RESULT 1	ID	APRE_PSEAE	STANDARD	PRT	432 AA
AC	003025				
DT	01-OCT-1993 (Rel. 27, Created)				
DT	16-OCT-2001 (Rel. 40, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	Alkaline protease secretion protein apre.				
GN	APRE OR PA1247.				
OS	Pseudomonas aeruginosa.				
OC	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;				
OX	Pseudomonas.				
NCBI_TaxID=287;					
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 15692 / PAOI;				
RX	MEDLINE=93051361; PubMed=1427098;				
RA	Duong F., Lazdunski A., Cami B., Mungier M.;				
RT	"Sequence of a cluster of genes controlling synthesis and secretion of alkaline protease in Pseudomonas aeruginosa: relationships to other secretory pathways."				
RL	Gene 121:47-54(1992).				
RM	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 15692 / PAOI;				
RX	MEDLINE=20437337; PubMed=10984043;				
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,				
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,				
RA	Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,				
RA	Brody L.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lam R.M.,				
RA	Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,				
RA	Reizer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;				
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an				
RT	opportunistic pathogen.";				
RL	Nature 406:959-964(2000).				
RM	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 15692 / PAOI;				
RX	MEDLINE=20437337; PubMed=10984043;				
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,				
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,				
RA	Garber R.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lam R.M.,				
RA	Brody L.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,				
RA	Reizer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;				
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an				
RT	opportunistic pathogen.";				
RL	Nature 406:959-964(2000).				
RM	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 15692 / PAOI;				
RX	MEDLINE=20437337; PubMed=10984043;				
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,				
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,				
RA	Garber R.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lam R.M.,				
RA	Brody L.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,				
RA	Reizer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;				
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an				
RT	opportunistic pathogen.";				
RL	Nature 406:959-964(2000).				
RM	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 15692 / PAOI;				
RX	MEDLINE=20437337; PubMed=10984043;				
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,				
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,				
RA	Garber R.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lam R.M.,				
RA	Brody L.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,				
RA	Reizer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;				
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an				
RT	opportunistic pathogen.";				
RL	Nature 406:959-964(2000).				
RM	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 15692 / PAOI;				
RX	MEDLINE=20437337; PubMed=10984043;				
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,				
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,				
RA	Garber R.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lam R.M.,				
RA	Brody L.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,				
RA	Reizer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;				
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an				
RT	opportunistic pathogen.";				
RL	Nature 406:959-964(2000).				
RM	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 15692 / PAOI;				
RX	MEDLINE=20437337; PubMed=10984043;				
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,				
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,				
RA	Garber R.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lam R.M.,				
RA	Brody L.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,				
RA	Reizer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;				
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an				
RT	opportunistic pathogen.";				
RL	Nature 406:959-964(2000).				
RM	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 15692 / PAOI;				
RX	MEDLINE=20437337; PubMed=10984043;				
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,				
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,				
RA	Garber R.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lam R.M.,				
RA	Brody L.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,				
RA	Reizer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;				
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an				
RT	opportunistic pathogen.";				
RL	Nature 406:959-964(2000).				
RM	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 15692 / PAOI;				
RX	MEDLINE=20437337; PubMed=10984043;				
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,				
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,				
RA	Garber R.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lam R.M.,				
RA	Brody L.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,				
RA	Reizer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;				
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an				
RT	opportunistic pathogen.";				
RL	Nature 406:959-964(2000).				
RM	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 15692 / PAOI;				
RX	MEDLINE=20437337; PubMed=10984043;				
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,				
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,				
RA	Garber R.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lam R.M.,				
RA	Brody L.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,				
RA	Reizer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;				
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an				
RT	opportunistic pathogen.";				
RL	Nature 406:959-964(2000).				
RM	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 15692 / PAOI;				
RX	MEDLINE=20437337; PubMed=10984043;				
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,				
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,				
RA	Garber R.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lam R.M.,				
RA	Brody L.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,				
RA	Reizer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;				
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an				
RT	opportunistic pathogen.";				
RL	Nature 406:959-964(2000).				
RM	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 15692 / PAOI;				
RX	MEDLINE=20437337; PubMed=10984043;				
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,				
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,				
RA	Garber R.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lam R.M.,				
RA	Brody L.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,				
RA	Reizer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;				
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an				
RT	opportunistic pathogen.";				
RL	Nature 406:959-964(2000).				
RM	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 15692 / PAOI;				
RX	MEDLINE=20437337; PubMed=10984043;				
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,				
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,				
RA	Garber R.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lam R.M.,				
RA	Brody L.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,				
RA	Reizer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;				
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an				
RT	opportunistic pathogen.";				
RL	Nature 406:959-964(2000).				
RM	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 15692 / PAOI;				
RX	MEDLINE=20437337; PubMed=10984043;				
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,				
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,				
RA	Garber R.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lam R.M.,				
RA	Brody L.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,				
RA	Reizer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;				
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an				
RT	opportunistic pathogen.";				
RL	Nature 406:959-964(2000).				
RM	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 15692 / PAOI;				
RX	MEDLINE=20437337; PubMed=10984043;				
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,				
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,				
RA	Garber R.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lam R.M.,				
RA	Brody L.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,				
RA	Reizer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;				
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an				
RT	opportunistic pathogen.";				
RL	Nature 406:959-964(2000).				
RM	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 15692 / PAOI;				
RX	MEDLINE=20437337; PubMed=10984043;				
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,				
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,				
RA	Garber R.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lam R.M.,				
RA	Brody L.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,				
RA	Reizer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;				
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an				
RT	opportunistic pathogen.";				
RL	Nature 406:959-964(2000).				
RM	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 15692 / PAOI;				
RX	MEDLINE=20437337; PubMed=10984043;				
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,				
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,				
RA	Garber R.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lam R.M.,				
RA	Brody L.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,				
RA	Reizer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;				
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an				
RT	opportunistic pathogen.";				
RL	Nature 406:959-964(2000).				
RM	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 15692 / PAOI;				
RX	MEDLINE=20437337; PubMed=10984043;				
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,				
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,				
RA	Garber R.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lam R.M.,				
RA	Brody L.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,				
RA	Reizer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;				
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an				
RT	opportunistic pathogen.";				
RL	Nature 406:959-964(2000).				
RM	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 15692 / PAOI;				
RX	MEDLINE=20437337; PubMed=10984043;				
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,				
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,				
RA	Garber R.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lam R.M.,				
RA	Brody L.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,				
RA	Reizer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;				
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an				
RT	opportunistic pathogen.";				
RL	Nature 406:959-964(2000).				
RM	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 15692 / PAOI;				
RX	MEDLINE=20437337; PubMed=10984043;				
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,				
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,				
RA	Garber R.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lam R.M.,				
RA	Brody L.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,				
RA	Reizer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;				
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an				
RT	opportunistic pathogen.";				
RL	Nature 406:959-964(2000).				
RM	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 15692 / PAOI;				
RX	MEDLINE=20437337; PubMed=10984043;				
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,				
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,				
RA	Garber R.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lam R.M.,				
RA	Brody L.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,				
RA	Reizer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;				
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an				
RT	opportunistic pathogen.";				
RL	Nature 406:959-964(2000).				
RM	[

FT DOMAIN 1 14 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 15 36 POTENTIAL.
 FT DOMAIN 37 432 PERIPLASMIC (POTENTIAL).
 FT CONFLICT 378 379 AL -> RV (IN REF. 1).
 SO SEQUENCE 432 AA; 48052 MW; 4C845C65DECF15EE CRC64;

Query Match 23.8%; Score 514.5; DB 1; Length 432;
 Best Local Similarity 31.6%; Pred. No. 1.9e-25;
 Matches 137; Conservative 88; Mismatches 182; Indels 27; Gaps 11;

QY 15 AAVARIGIATLVGLGMAFAFLDS-AVIANGVSAEVSODVQHLLEGMLKILVR 72
 D 11 AAVARIGIATLVGLGMAFAFLDS-AVIANGVSAEVSODVQHLLEGMLKILVR 69
 QY 73 EGEKAKAGVLELDPTQANAAAGITRNQYALKAMEARLLAERDQPSISFPADLTQOR 132
 D 70 DGOHVEAGEPILIRMEPTQARNVDSLNRVYANARLQARLQAEYDGRRTLEGPAQLAQDA 129
 QY 133 ADPMVAAIADDOAQFERRROT-IGQVYDLNMAQRLQYQSELEGIDROTQGL- LKQDLG 188
 D 130 PLPLTIGERL-ELQRLHLSROTALANESALRANIEGRRAQLEGL-ROTEGNORLOQRL- 186
 QY 189 FIEDELIDRLKLYDKGLVPRRLALE-----ARAGSLSGSIGRLTADRKAQVQASD 241
 D 187 -LNSQISGARDLAEGYVPRNQLLEQERQALAEVNARLSESSQRFQIR-----QSLAE 238
 QY 242 TOLKVRQIKOEFEEDVQSITETRVRLAEVTEKEVYASDAQRIKIVSPVQCTAQNREF 301
 D 239 AOMRIAQKEEYERKEVNGQLAETQVNAETLWEELSSARYELRHAELIRAPVSGYVAGLVKF 298
 QY 302 TEGAVVRAAEPLVDIAPEDEAFVIOAHQPTDNDVNHGKMTVEVLPFHSAGNDPBRH 361
 D 299 TDGAVIGGELMTIVPRSDLEVEGQALVNLVDRHSGLEVEMLFTFNSKIPRVTGE 358
 QY 362 DPVAVADRISDPQOARLEIGIVRVYKQRLPHLG-RVTAGMPAQVIVPTGERTVQYL 420
 D 359 VTWVSADRLDEONKQPYALRAQVDAAM-GKLKGLDIRGMAVQVEVFRGERSLNYL 417
 QY 421 FSPRLDILRTTMR 434
 D 418 FKPLPRAHVALAE 431

RESULT 2

PRTE_ERWCH STANDARD; PRT; 448 AA.
 AC P23597;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE Proteases secretion protein prte.
 GN PRTE.
 OS *Erwinia chrysanthemi*.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Pectobacterium.
 OX NCBI_Taxid=556;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90228333; Pubmed=2184029;
 RA Letoite S., Delepelate P., Wandersman C.;
 RT "Protease secretion by *Erwinia chrysanthemi*: the specific secretion
 RT functions are analogous to those of *Escherichia coli*
 RT alpha-hemolysin".
 RL EMBL J. 9:1375-1382(1990).
 CC -1- FUNCTION: INVOLVED IN THE SECRETION OF PROTEASES A, B, C AND G.
 CC -1- SUBCELLULAR LOCATION: INNER MEMBRANE-BOUND.
 CC -1- SIMILARITY: BELONGS TO THE HLYD FAMILY OF SECRETION PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X53253; CAAB37343.1;
 DR EMBL; M60395; AAA63635.1;
 DR PIR; S12526;
 DR InterPro; IPR002215; HLYD.
 DR InterPro; IPR003997; RCD.
 DR Pfam; PF00529; HLYD.1.
 DR PRINTS; PR01490; RTXTXIND.
 DR PROSITE; PS00543; HLYD_FAMILY.1.
 KW Transmembrane; Inner membrane; Transport.
 FT DOMAIN 1 30 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 31 50 POTENTIAL.
 FT DOMAIN 51 448 PERIPLASMIC (POTENTIAL).
 SO SEQUENCE 448 AA; 49579 MW; 5112ACE1AFB84473 CRC64;

Query Match 23.4%; Score 506.5; DB 1; Length 448;
 Best Local Similarity 28.8%; Pred. No. 6.2e-25;
 Matches 128; Conservative 93; Mismatches 191; Indels 33; Gaps 7;

QY 8 PPTDQFQAVARIGYGIATLVGLGMAFAFLDSAVIANG-VSAEVSODVQHLLEGML 66
 D 20 RASRDEERALLGKMWLVAGFGFLMALLAPLDKGVAVQGNVYSGNRKVIQHMGQIV 79
 QY 67 AKILVGEKVKAGVLELDPTQANAAAGITRNQYALKAMEARLLAERDQPSISFPA 126
 D 80 DRIOVKDGDVRAAGQVLLTLNVAVARTTSEGLSQYDOLIRAREARLLAERQNSLIATP 139
 QY 127 DLTQSRADPMVAAIADDOAQFERRROTIGQVYDLNMAQRLQYQSELEGIDROTQGLKQ 186
 D 140 RLQARQRPMAIITLQD-----DLNRQOOSKLEIDVQASIDGLTTS 185
 QY 187 LGFTE-----DELIDRLKLYDKGLVPRRLALEARAGSLSGSIGRLTADR 232
 D 186 LGALQKVMSSKQSEGLTQDLQRLPLAADNVPYPRNMLETERRLPAYSELAQTSEV 245
 QY 223 SKAVOGASDTOLKVRQIKOEFEEDVQSITETRVRLAEVTEKEVYASDAQRIKIVSPV 292
 D 246 GRTDRDIOQKRLIAQROOEYDEKVENSLSDVQKLNELVSIQREKADFNLANVQVAPVA 305
 QY 293 GTAQNLREFTEGAVVRAAEPLVDIAPEDEAFVIOAHQPTDNDVNHGKMTVEVLPFHS 352
 D 306 GTVYDMKLTETGVIYARQVQMDVPEQDPLLVGRIRPEVAVQVWSGLDFTLASQ 365
 QY 353 AGNDPERRHDPVAVADRISDPQOARLEIGIVRVYKQRLPHG-LRG-RVTAGMPAQVIVP 410
 D 366 STTFRVPGTVLLSADRLVD-EKDGTPYGL-RIOVSEGRSLHGLEIKGMVQGVFR 423
 QY 411 TGERTVQYLFSPRLDILRTTMR 435
 D 424 TGRSFIVLYFKPLMDRMHIALTEE 448

RESULT 3

RT3D_ACTPL STANDARD; PRT; 477 AA.
 AC Q08633;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE RTX-III toxin determinant D (ApX-III) (Cytolysin IIID) (CLY-IIID)
 DE (Toxin RTX-III secretion protein D).
 GN APXIIID OR CLYIIID OR RTX.
 OS *Actinobacillus pleuropneumoniae* (Haemophilus pleuropneumoniae).
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC *Actinobacillus*.
 OX NCBI_Taxid=715;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Serotype 2;
 RX MEDLINE=93263992; Pubmed=8494611;
 RA Chang Y.-F., Shi J., Ma D.-P., Shin S.-J., Levin D.H.;


```

RT      "Molecular analysis of the Actinobacillus pleuropneumoniae RTX
RT      toxin-III gene cluster.";
RL      DNA Cell Biol. 12:351-362(1993).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN-405 / Serotype 8;
RX      MEDLINE-95012630; PubMed-7927703;
RA      Jansen R., Briatore J., van Geel A.B.M., Kamp E.M., Gielkens A.L.J.,
RA      Smits M.A.;
RT      "Genetic map of the Actinobacillus pleuropneumoniae RTX-toxin (ApX)
RT      operon: characterization of the ApXIII operons.";
RL      Infect. Immun. 62:4411-4418(1994).
CC      -1- FUNCTION: INVOLVED IN THE TRANSPORT OF THE TOXIN RTX-III.
CC      -1- SUBCELLULAR LOCATION: Inner membrane-bound (POTENTIAL).
CC      -1- MISCELLANEOUS: THE SEQUENCE SHOWN IS THAT OF SEROTYPE 2.
CC      -1- SIMILARITY: BELONGS TO THE HLYD FAMILY OF SECRETION PROTEINS.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
-----
DR      EMBL, L12145; AAA21926.1; -.
DR      EMBL, X80055; CAA56360.1; -.
DR      InterPro: IPR002215; HlyD.
DR      InterPro: IPR003997; RtxD.
DR      Pfam: PF00529; HlyD; 1.
DR      PRINTS: PR01490; RTXTOXIND.
DR      PROSITE: PS00543; HLYD_FAMILY; 1.
KM      Cytolysis; Transport; Transmembrane; Inner membrane.
FT      DOMAIN 1 59 CYTOPLASMIC (POTENTIAL).
FT      TRANSMEM 60 77 POTENTIAL.
FT      DOMAIN 78 477 PERIPLASMIC (POTENTIAL).
FT      VARIANT 44 44 R -> A (IN SEROTYPE 8).
FT      VARIANT 132 132 C -> A (IN SEROTYPE 8).
FT      VARIANT 185 186 GA -> NT (IN SEROTYPE 8).
FT      VARIANT 376 376 V -> I (IN SEROTYPE 8).
SQ      SEQUENCE 477 AA; 54780 MW; D229FICABD7E90A CRC64;

Query Match 14.5%; Score 313.5; DB 1; Length 477;
Best Local Similarity 25.3%; Pred. No. 9e-13;
Matches 112; Conservative 88; Mismatches 190; Indels 53; Gaps 13;

OY      19 ICGIITLFTVGLGMAAPLDSAVIANG-VYSAEVSODVCHLEGMAKLIVRGCEVY 77
DB      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      60 IAYLIMLFLELAIV-ISIISKVEIYASATGKLVFSCHSKEIKRIENALVYKDGOPY 118
OY      78 KAGOVLEL-----DPTQANNAAGITRNOVALKAMEARLLAERDQPSISF-PADLT 129
DB      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      119 EKGOLLINTALGCDADKOKTVSLGLELDGTRYKSLYKST--ENNRPLPLDFNOADPD 176
OY      130 S-QRADPNVYARAIDE-----QAQFERRQTIQGVDMNARLOYOSELEGIDR 178
DB      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      177 SVOEDKGTGARHLITEQFTWQOKYOKELAYORKQAEKQYVLANIRKYESASRIE---- 232
OY      179 QYQGLKDLQGFIEDELIDRLKLYDKGLVPRPLLALEAR---AGLSSTIGRLTRDRSK 234
DB      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      233 -----KERISDRLKLYDVKSISKHLLAENRYEASNELSYQSHLKEVED 280
OY      235 AVOGASDPLQAKYROIKEFEFEVOSITPTTRVRLAEVTEKEVYVSAQOKRIK--IYSPVN 292
DB      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      281 LTKAEDLKLVTQLFKSDLEKTLQOVIQREKQLTLELENE-----QOKLASITRAPVS 334
OY      293 GTAQNLRFETEGAVYRAEPLVDIAPEDEAFYIOAHFOPTVDVNMHMGVETVRLPAFVS 352
DB      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      335 GYVQQLKTKTGCGVYTTAETLAVIAPEDDVEVLSALIQKDKDGFGEIAGEANIKYETFFY 394
OY      353 AGNPERRHDPAVADRISDPQOKARLGLIVRVKOLPPLRG-RTYAGPAQVIVPT 411
DB      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      395 TRYGYLYGKVKTTITLDAIEHPQ-LGLVFNSTIEINKKTLTDGDKETQLSGSMVIAEIKT 453

```

```

OY      412 GERTVQVLESPLDRLRTTMR 434
DB      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      454 GERSVIFLSPLSESTESLRE 476

RESULT 4
ID      HLY4_ECOLI STANDARD; PRT; 478 AA.
AC      P09386;
DT      01-MAR-1989 (Rel. 10, Created)
DT      01-MAR-1989 (Rel. 10, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Hemolysin secretion protein D, chromosomal.
GN      HLYD.
OS      Escherichia coli.
OC      Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC      Escherichia.
OX      NCBI_Taxid=562;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-196 / Serotype O4;
RX      MEDLINE-85234404; PubMed-3891743;
RA      Fellmeier T., Pellet S., Welch R.A.;
RT      "Nucleotide sequence of an Escherichia coli chromosomal hemolysin.";
RL      J. Bacteriol. 163:94-105(1985).
RN      [2]
RP      TOPOLOGY.
RX      MEDLINE-92357011; PubMed-1495479;
RA      Schuelein R., Gentschev I., Mollenkopf H.-J., Goebel W.;
RT      "A topological model for the haemolysin translocator protein HlyD.";
RL      Mol. Gen. Genet. 234:155-163(1992).
CC      -1- FUNCTION: INVOLVED IN THE TRANSPORT OF HEMOLYSIN A.
CC      -1- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane.
CC      -1- SIMILARITY: BELONGS TO THE HLYD FAMILY OF SECRETION PROTEINS.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
-----
DR      EMBL, M0133; AAA23977.1; -.
DR      PIR: D24433; LEED.
DR      InterPro: IPR002215; HlyD.
DR      InterPro: IPR003997; RtxD.
DR      Pfam: PF00529; HlyD; 1.
DR      PRINTS: PR01490; RTXTOXIND.
DR      PROSITE: PS00543; HLYD_FAMILY; 1.
KM      Hemolysis; Transport; Transmembrane; Inner membrane; Signal-anchor.
FT      DOMAIN 1 59 CYTOPLASMIC (PROBABLE).
FT      TRANSMEM 60 80 SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)
FT      DOMAIN 81 478 PERIPLASMIC (PROBABLE).
SQ      SEQUENCE 478 AA; 54591 MW; 9E9EDC42EC8D8089 CRC64;

Query Match 13.3%; Score 288.5; DB 1; Length 478;
Best Local Similarity 25.2%; Pred. No. 3.4e-11;
Matches 116; Conservative 94; Mismatches 192; Indels 59; Gaps 18;

OY      1 MKPRKQRTDNRQANRIGYIATFY-GILGMAAFAPLDSAVIANGVYV-AEVSQDV 58
DB      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      49 IETPVSRP-----RLVAYTFMGFLVAFILSYLG-----QVELVATANKRLTLGSKSEI 99
OY      59 OHLEGMAKLIVREGEKAKAGVLELDPTQANNAAGITRNOVALK-----MEARLLA 114
DB      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      100 KPIENSTIVAEITVKEBSRKQDVILKT--TALGAEADPLTKQSSILQARLQIRQIUIS 157
OY      115 ---ERDQPSISFPADLTQSR-ADPNVYARAIDEQAQF--TERRQTIQGVLD--MNAQRL 167
DB      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      158 RSEIENKLPKLPELPDEPPYQONVSEEVRLTLRIKQFSTWQONOKYOKELNIDKKRAERL 217

```

QY 168 QYSEIESIDPOTGLDQGFIEDELIDRLKLYDKGLVPRRLALEAR---AGSLSG 223
 Db 218 TILRLNKYEVSVREKSRDL-----DFRSLHKQAIKHAHVEQENKYEANELRV 270
 QY 224 SIGRLDRAKAVGASPTOLKVRQIKQEFPEQVYSQSTETERRVLAETKEVEVASDQK 283
 Db 271 YKSOEQIESLTSKKEFYQLVTOLEFKNEIDKRLQYTDSTELLTELEKNE---ERQK 326
 QY 284 RIKIVSPVNGTAQNLRFETEGAVVRAAEPLVDIAPEDEAFYIOAHFOPTDVNVMGMVVT 343
 Db 327 ASVIRAPVSGKVKQQLKVTBEGVVTATLTAIVYEDDTLEVTAIVQKQDGLGFIWQONA 386
 QY 344 EVRLPAFHSAGNPDPERRHDPAVA-----DRISDPQKQARLFLGIVRVYKOLP---P 393
 Db 387 IIRKEAF-----PYTRYGLVCKVKNNINDAIED-OKGLGVFNVIYSVEENDLSTGK 438
 QY 394 HLRGRVTAAGMPAQYIVPTGERTVQLYFSPRLDRLRTMRE 434
 Db 439 HI--PLSSGMVAVTAIRKTMRSVISYLLSPLEESVTESLHE 477

RESULT 5

HLVD_ACTAC STANDARD: PRT: 477 AA.

AC P18790;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Leukotoxin secretion protein D.
 GN LKTD OR ALTD.
 OS Actinobacillus actinomycetemcomitans (Haemophilus actinomycetemcomitans).
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae; Actinobacillus.
 OX NCBI_Taxid=714;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-JF2;
 RX MEDLINE=90384843; PubMed=2402458;
 RA Guthmiller J.M., Craig E., Cagle M.P., Kolodrubetz D.;
 RT "Sequence of the lktD gene from Actinobacillus actinomycetemcomitans";
 RL Nucleic Acids Res. 18:5292-5292(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92072446; PubMed=1961107;
 RA Lally E.T., Golub E.E., Kieba I.R., Taichman N.S., Decker S., Berthold P., Gibson C.W., Demuth D.R., Rosenblum J.;
 RT "Structure and function of the B and D genes of the Actinobacillus actinomycetemcomitans leukotoxin complex";
 RL Microb. Pathog. 11:111-121(1991).
 CC -1- FUNCTION: INVOLVED IN THE TRANSPORT OF THE HEMOLYSIN/LEUKOTOXIN.
 CC -1- SUBCELLULAR LOCATION: Inner membrane-bound (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE HLVD FAMILY OF SECRETION PROTEINS.
 CC -1- SIMILARITY: THE N-TERMINAL (AA 1-153) IS SIMILAR TO E.FAECALIS PORE FORMING PROTEIN EBSA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X53956; CA37907.1;
 DR PIR: S11315; S11315.
 DR PIR: B61378; B61378.
 DR InterPro: IPR002215; HLVD.
 DR InterPro: IPR003997; RLXD.
 DR Pfam: PF00529; HLVD; 1.
 DR PRINTS: PR01490; RTXTOXIN.D.

DR PROSITE; PS00543; HLVD_FAMILY; 1
 KW Hemolysis; Cytolysis; Transport; Transmembrane; Inner membrane.
 FT DOMAIN 1 59 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 60 80 POTENTIAL.
 FT DOMAIN 81 477 PERIPLASMIC (POTENTIAL).
 SQ SEQUENCE 477 AA; 54651 MW; 54F20128C4D6260 CRC64;

Query Match 12.4%; Score 267.5; DB 1; Length 477;
 Best Local Similarity 23.7%; Pred. No. 7e-10;
 Matches 108; Conservative 87; Mismatches 207; Indels 53; Gaps 15;

QY 6 IORPDNFQAVARIGYGIATLTPVGLGMAFAPLDSAVINMGVVS-AEVSQDVQHEGG 64
 Db 49 IETPVSN--APRFVSYSLMFLTLAII-VSIFSNVEIATASGRFALSGRKEIKPIENS 105
 QY 65 MLAKLIVREGKAVAGVQLFELDPDQANAAGIRNQVVALKAEARL-----LAERDQ 119
 Db 106 LKHLIFAVEGEYVKKGLLKLTLALGAEADYLLKRTSLSQAKLEEFYKSLLEAVEKQDL 165
 QY 120 PSISE-PADLTSSORADPWAVARAIADEQAQTERQTOGGVYL--MNAORLYOSEIEGI 176
 Db 166 PILDFSKIDLPFMTENDQKRRTLLIEQFSTWQKQKHQKTLNLKKKEKSTYLARI--- 222
 QY 177 DROTGLKQDQGFIDE---LIDRLKLYDKGLVPRRLALEARAGSLSGSIGRLTDRS 233
 Db 223 -----KKYEGLINTEGVRLDDERFALYKEHAIKHVLDEE---NKYQDAINLEEVYKA 272
 QY 224 KAVQ-----GASPTOLKVRQIKQEFPEQVYSQSTETERRVLAETKEVEVASDQKRIK 286
 Db 273 SLMOYENEVLLAKEQELVTLDFKNDILDKKQATDNVNLITFELDKN---NORQVSE 328
 QY 287 IVSPVNGTAQNLRFETEGAVVRAAEPLVDIAPEDEAFYIOAHFOPTDVNVMGMVTEVR 346
 Db 329 IRAPVSGTQOLKVTHTIDGVVTTAETLTMVYVPEBDSLEVTALQNKDGLGFEQGEQEVYIK 388
 QY 347 LPAF-----HSAGNPDPERRHDPAVAVDRISDPQKQARLFLGIVRVYKOLPRLRG-RV 399
 Db 389 VEAPPYTRYGLTGKVN-----ITLDAIEHP-KLGLVFNTIELDKRTLSTEKEIPL 441
 QY 400 TAGMPAQYIVPTGERTVQLYFSPRLDRLRTMRE 434
 Db 442 SAGMEITAEIKTGRSVISYLLSPLEESIDKSLAE 476

RESULT 6

HLVD_ECOLI STANDARD: PRT: 478 AA.

AC P06739;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hemolysin secretion protein D, plasmid.
 GN HLVD.
 OS Escherichia coli.
 OC Plasmid Inc12 PHU152.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.
 OX NCBI_Taxid=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC HESS J., WEIS W., VOGEL M., GOEBEL W.;
 RT "Nucleotide sequence of a plasmid-encoded hemolysin determinant and its comparison with a corresponding chromosomal hemolysin sequence";
 RL FEWS Microbiol. Lett. 34:1-11(1986).
 RN [2]
 RP TOPOLOGY.
 RX MEDLINE=92357011; PubMed=1495479;
 RA Schuelein R., Gentschew I., Mollenkopf H.-J., Goebel W.;
 RT "A topological model for the haemolysin translocator protein HlyD";
 RL Mol. Gen. Genet. 234:155-163(1992).
 CC -1- FUNCTION: INVOLVED IN THE TRANSPORT OF HEMOLYSIN A.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane.
 CC -1- SIMILARITY: BELONGS TO THE HLVD FAMILY OF SECRETION PROTEINS.

Query Match	Best Local Similarity	Score	DB	Length
12.3%	26.2%	Pred. No. 7.9e-10;		
Matches 110;	Conservative 71;	Mismatches 181;	Indels 58;	Gaps 16;
Sequence 440 AA;	48011 MM;	B786328A92DD0D48 CRC64;		
NCBI_TaxID=520;				
SEQUENCE FROM N.A.				
STRAIN-18323;				
MEDLINE=89091151;				
PubMed=2905265;				
Glaser P., Sakamoto H., Bellalou J., Ullmann A., Danchin A.;				
"Secretion of cyclolysin, the calmodulin-sensitive adenylyl cyclase-				
haemolysin bifunctional protein of Bordetella pertussis.";				
EMBO J. 7:3997-4004(1988)				
- FUNCTION: CYAD IS NECESSARY FOR TRANSPORT OF CALMODULIN-SENSITIVE				
ADENYLYL CYCLASE-HEMOLYSIN (CYCLOLYSIN).				
- SUBCELLULAR LOCATION: Inner membrane-bound (POTENTIAL).				
- SIMILARITY: BELONGS TO THE HYD FAMILY OF SECRETION PROTEINS.				
This SWISS-PROT entry is copyright. It is produced through a collaboration				
between the Swiss Institute of Bioinformatics and the EMBL outstation -				
the European Bioinformatics Institute. There are no restrictions on its				
use by non-profit institutions as long as its content is in no way				
modified and this statement is not removed, usage by and for commercial				
entities requires a license agreement (See http://www.isb.ch/announce/				
or send an email to license@isb.slb.ch).				
EMBL; X14199; CAA32413.1; -				
PIR; S02387; BVRBCD.				
InterPro; IPR002215; HYD.				
InterPro; IPR003997; RKXD.				
Pfam; PF00529; HYD. 1.				
PRINTS; PR01490; RXTXOIND.				
PROSITE; PS00543; HYD_FAMILY. 1.				
Hemolysin; Transport; Transmembrane; Inner membrane.				
DOMAIN 1 55 CYTOPLASMIC (POTENTIAL).				
TRANSMEM 56 75 PERIPLASMIC (POTENTIAL).				
DOMAIN 76 440 POTENTIAL.				
SEQUENCE 440 AA; 48011 MM; B786328A92DD0D48 CRC64;				
16 VARIETYIATLFTVYLLGMAAFAPLDSAVIANG-VVSAEVSODVYHLEGMLAKILVREG 74				
57 LARI---LMALA-AGALVWSVVGKREIYVHAAGKVVPVGGOSKILIASSEGGRVAVRLVADN 112				
75 EKVKAQVYLFELDPQVNAAGSITRNQVYALKAMEARLAEFDQRPISFPADLTSGRAD 134				
113 SRVAAGDVLRLD-----AGVTAERKRWVQAQ--ARODEARSRAMIRALDTGRA- 162				
135 PMVAATIADEQAQFERRQTIQGVUDLMNAQROLOVSEIEGIDRQTQGLKQDIEDEL 194				
163 PVALALPAPDPPGMAAO-----SYDSQIADYQAOQRSTIEAIAITTYRDVGLVTOIA 213				
195 IDLRLLYDKGLVPRRLALAEAGSLSGISGRILTRADRSKAVOGASDPTQVKVRIKQEFF 254				
214. HAHRELRRDGDVSGQAVYLEKEQARLTLEGR-----RQSEAGRAALQVOTR---ROAF- 263				
255 EGVOSQIETRY---RLAEVTEKEVNVASDAQ-RKIIVSPVNGTQONLRPFPEGAVVRAA 310				
264 -----ETVLARKLQAQEQEIAKTSQAORSRSLVTLPAVDGVQVQVVALTEGTAIVAT 315				
311 EPLVDAIEDEAFVIAHFQPTDVNVHMGWTEVRLPAFSAGNPDPERHDPPVAVADRI 370				
316 QPLMNVVSGAGIQVQAOQDLSNDIGFVAGAPATYKGAIDYTKTGKILEGKVLVYSPPTV 375				
371 SDPOKO--ARLEFLGI---VRVDVKQLPRLHGRVTAQMPAOVIVTGERVLYQLESPL 424				
376 VDDRQOQHRYVTIALAHPALEVDGK---PRL---LKEGMVAQADIRTSRRILEYLLSPV 429				

DT 01-OCT-1996 (Rel. 34, last sequence update)
 DT 15-JUN-2002 (Rel. 41, last annotation update)
 DE Leukotoxin secretion protein D.
 GN LKTD.
 OS Pasteurella haemolytica-like sp. (strain 5943b).
 CC Bacteria: Proteobacteria; gamma subdivision: Pasteurellaceae;
 CC Pasteurella.
 CC NCBI_TaxID=28165;
 RX MEDLINE=93239320; PubMed=8478098;
 RA Chang Y.-F., Ma D.-P., Shi J., Chengappa M.M.;
 RT "Molecular characterization of a leukotoxin gene from a Pasteurella
 RT haemolytica-like organism, encoding a new member of the RTX toxin
 RT family.";
 RL Infect. Immun. 61:2089-2095(1993).
 CC -1- FUNCTION: INVOLVED IN THE TRANSPORT OF THE LEUKOTOXIN.
 CC -1- SUBCELLULAR LOCATION: Inner membrane-bound (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE HLXD FAMILY OF SECRETION PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L12148; AAA16446.1;
 DR InterPro: IPR002215; HLXD.
 DR InterPro: IPR003997; RTX.
 DR Pfam: PF00529; HLXD.1
 DR PRINTS: PR01490; RTXTOXIND.
 DR PROSITE: PS00543; HLXD_FAMILY.1
 KW Hemolysis; Cytolysis; Transport; Transmembrane; Inner membrane.
 FT DOMAIN 1 77 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 78 98 POTENTIAL.
 FT DOMAIN 99 478 PERIPLASMIC (POTENTIAL).
 SO SEQUENCE 478 AA; 54650 MW; 221461a69482913a CRC64;
 Query Match 11.2%; Score 243.5; DB 1; Length 478;
 Best Local Similarity 24.4%; Pred. No. 2.3e-08;
 Matches 113; Conservative 73; Mismatches 174; Indels 103; Gaps 18;
 QY 31 LIGNAAPPLDSAVIANCVSAEV-----SODVHLEGMLATIREGEKV 77
 DB 59 LIALIMLFALVAIVLASVSKVEIVATAPGKLTFSGRSKETPIENTIVOEIFVYDCGFV 118
 QY 78 KAGOVLEFLDP--TOANAAAGITRNQYVALKAMEARLL--AERDQRSISFPADLT-- 129
 DB 119 EKGGLVSLTRALGSDADIKTTITSLAKLENRYOTLLTAIEKESLPVY---DLNTE 174
 QY 130 ---SORADPWVAAIADP-----QAQFTER-----RDTIOGVYDLNNAORLOYOS 171
 DB 175 FKDSSEEDRLIKHLIEQVYTWQKOKTKLAVKRKAQKAEQITSA-----YVR 223
 QY 172 EIEGIDQTQGLKQQLGFIETDELIDKLYDKGLVPRRLLEARRAGSLSGIGRLTAD 231
 DB 224 KYEGATRIEQQ-----EKLKDFRLTXQOKSLSKHELLSOENKAIEAO---NELAVY 270
 QY 232 RSKAVOGASD-----TOLKVRQIKOEFEFOVSOSITETVRVLAEVEKEVASDQAKR 284
 DB 271 RSKLNELESLDLNKEELLETTOFFKSDVLEKLOHLENEHOLEL-EKN---NQRQA 326
 QY 285 IKIVSPVNGTAQNLRFFTEGAIVVRAAEPLVDIAEDFAVQAHFQPTDVNVHMGWTE 344
 DB 327 SMIAPIASGVYQQLKIHIGGVVTAETIMTIYVEDDVELEFALIQNDIGFVAAQGEVI 386
 QY 345 VRLPAFHSAGNPDERHPDVA---VADRISS---DPQKQARL---FLGIVRVYVKOL- 391
 DB 387 IKVETP-----PYTRYGYITGRIKIHISPADEQGNLGVFNATVYSIDKQALSS 434
 QY 392 PPHLRGRTVAGMPAQVIVPGERIVLOYLSPDLDTLTATRE 434

DB 435 PDGKHIEGLGMLTTAEIKTGERSVSYLSPLSESVYSURE 477
 RESULT 9
 RTID:ACTPL STANDARD; PRT: 477 AA.
 ID RTID:ACTPL
 AC P26761;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, last sequence update)
 DT 15-JUN-2002 (Rel. 41, last annotation update)
 DE RTX-I toxin determinant D (APX-ID) (HLX-ID) (Cytolysin ID) (CLR-ID)
 DE (toxin RTX-I secretion protein D).
 GN APXID OR CLYD OR HLXD OR APPD.
 OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
 CC Bacteria; Proteobacteria; gamma subdivision: Pasteurellaceae;
 CC Actinobacillus.
 CC NCBI_TaxID=715;
 RX MEDLINE=91317735; PubMed=1860823;
 RA Chang Y.-F., Young R., Struck D.K.;
 RT "The Actinobacillus pleuropneumoniae hemolysin determinant: unlinked
 RT appca and appbd loci flanked by pseudogenes.";
 RL J. Bacteriol. 173:5151-5158(1991).
 RN [1]
 RP SEQUENCE FROM N.A.
 RP STRAIN-Serotype 5;
 RC MEDLINE=92040145; PubMed=1937809;
 RX Smiths M.A., Briatore J., Jensen R., Smith H.E., Kamp E.M.,
 RA Gielkens A.L.;
 RT "Cytolysins of Actinobacillus pleuropneumoniae serotype 9.";
 RL Infect. Immun. 59:4497-4504(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RP STRAIN-S 4074 / Serotype 1;
 RC MEDLINE=94237497; PubMed=8181764;
 RA Frey J., Haldemann A., Nicolet J., Boffini A., Prentki P.;
 RT "Sequence analysis and transcription of the apxi operon (hemolysin I)
 RT from Actinobacillus pleuropneumoniae.";
 RL Gene 142:97-102(1994).
 CC -1- FUNCTION: INVOLVED IN THE TRANSPORT OF THE TOXIN RTX-I AS WELL AS
 CC THAT OF RTX-II.
 CC -1- SUBCELLULAR LOCATION: Inner membrane-bound (POTENTIAL).
 CC -1- MISCELLANEOUS: THE SEQUENCE SHOWN IS THAT OF SEROTYPE 5.
 CC -1- SIMILARITY: BELONGS TO THE HLXD FAMILY OF SECRETION PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M65808; AAB00967.1;
 DR EMBL: X61112; CAA43426.1;
 DR EMBL: X68595; CAA48588.1;
 DR PIR: B40366; B40366.
 DR PIR: S18856; S18856.
 DR InterPro: IPR002215; HLXD.
 DR InterPro: IPR003997; RTX.
 DR Pfam: PF00529; HLXD.1.
 DR PRINTS: PR01490; RTXTOXIND.
 DR PROSITE: PS00543; HLXD_FAMILY.1.
 KW Hemolysis; Cytolysis; Transport; Transmembrane; Inner membrane.
 FT DOMAIN 1 59 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 60 80 POTENTIAL.
 FT DOMAIN 81 477 PERIPLASMIC (POTENTIAL).
 FT VARIANT 14 14 A -> R (IN SEROTYPES 1 AND 9).
 FT VARIANT 94 94 G -> D (IN SEROTYPES 1 AND 9).
 FT VARIANT 113 113 E -> Q (IN SEROTYPES 1 AND 9).
 FT VARIANT

FT VARIANT 228 228 S -> I (IN SEROTYPES 1 AND 9).
 FT VARIANT 335 335 Y -> GT (IN SEROTYPES 1 AND 9).
 SQ SEQUENCE 477 AA; 54727 MW; B841D87AF7DAD4C2 CRC64;
 Query Match 10.8%; Score 233.5; DB 1; Length 477;
 Best Local Similarity 22.9%; Pred. No. 9.6e-08;
 Matches 105; Conservative 85; Mismatches 185; Indels 83; Gaps 17;
 19 IGYGIATFVGLGMAFAPLDSAVIANGVVS-AEVSODVOHLEGGMLAKTIVEGKV 77
 Db IAYLMFLFALV-ISTVSHVEIVATGKLAESGRSKETPIEMALVKELFVEDGQFV 118
 QY 78 KAGVLFELDPTQANAAGITRNQYVALKAMEAR--LLAE--RPDRPSISPPAD--LTS 130
 Db 119 EKDOLLHLTALGADADQDKRSSLSLTKLERYREITILEAVADRLLIELTKDEFKHA 178
 QY 131 QRADPVVARAIADDAQFTER-----RQTIQGVDDLMNQRLQYSEIEGIDRQT 180
 Db 179 TEEDKTRIRFLITEQFEAMQKQYKELALORREAKQTVLANIRKYGSSRVE----- 232
 QY 181 QGLKDQGLFIDELELDLRLKLDKGLVPRRLALEAR---AGLSGSGITRLADRSKAV 236
 Db 233 -----NRLKDLKLFNSSTSKHDVLTQENRHIENAVNELAVYKSRINEVESDLR 282
 QY 237 QGASDTOLKVRQIKOEFEQVQSITETRVRLAEVTEKEVVASDAOKRIKIV--SPVNGT 294
 Db 283 QAKEIHLITOLFADILEKLQNV-EAEKQSLSELEK-----NQRQIASYTRAPVS-Y 335
 QY 295 AQNTREFTEGAVVRAEPLVDIAPDEAFVIOAHFOPTVDNVHMGVTEVRLPAFHSAG 354
 Db 336 VQQLKTHVGAVVTAETLVIATAPEDVDLEVLTALIQNDIGFIEVGQAVIKVEF----- 391
 QY 355 NPDEPRHDPA-----VADRI SDPQKARLFLGIVRVVVKQPLRLR----- 398
 Db 392 -----PYTRYGYLMKRVKNITLLEALEHPQ-LGLVENSISIDRKT---LSGDKGE 438
 QY 399 --VTAGMPAQVIPTGERTVLOYLFSPRLDRLRTTMR 434
 Db 439 IELSGMSVTAETIKTGRSVISYLLSPLEESVSESLRE 476
 RESULT 10
 HLVD_PASHA STANDARD: PRT: 478 AA.
 AC P16534;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Leukotoxin secretion protein D.
 GN LKTD.
 OS Pasteurella haemolytica.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Mannheimia.
 OX NCBI_TaxID=75985;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Serotype A1;
 RX MEDLINE=891213172; PubMed=2914876;
 RA Strathdee C.A., Lo R.Y.C.;
 RT "Cloning, nucleotide sequence, and characterization of genes encoding
 the secretion function of the Pasteurella haemolytica leukotoxin
 determinant";
 RT J. Bacteriol. 171:916-928(1989).
 RL [3]
 RP SEQUENCE FROM N.A.

RC STRAIN-Serotype A1 / PHL101;
 RA MEDLINE=89210283; PubMed=2707120;
 RX Highlander S.K., Chidambaram M., Engler M.J., Weinstein G.M.;
 RT "DNA sequence of the Pasteurella haemolytica leukotoxin gene
 cluster";
 RL DNA 8:15-28(1989).
 CC -1- FUNCTION: INVOLVED IN THE TRANSPORT OF THE LEUKOTOXIN.
 CC -1- SUBCELLULAR LOCATION: Inner membrane-bound (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE HLVD FAMILY OF SECRETION PROTEINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL: M20730; AAA25531.1;
 CC EMBL: M24197; AAA25545.1;
 DR PIR: B32051; B32051.
 DR PIR: S29518; S29518.
 DR InterPro: IPR002215; HLVD.
 DR InterPro: IPR003997; RCD.
 DR Pfam: PF00529; HLVD; 1.
 DR PRINTS: PR01490; RITYTOXIND.
 DR PROSITE: PS00543; HLVD_FAMILY. 1.
 KW Hemolysis; Cytolysis; Transport; Transmembrane; Inner membrane.
 FT DOMAIN 1
 FT TRANSMEM 60 80
 FT DOMAIN 81 478
 FT POTENTIAL 59
 FT PERIPALMIC (POTENTIAL).
 FT CONFLICT 18 18
 FT CONFLICT 47 47
 FT CONFLICT 150 150
 FT CONFLICT 236 236
 FT CONFLICT 267 267
 SQ SEQUENCE 478 AA; 54761 MW; E2B8B5F5AF988C1 CRC64;
 Query Match 10.8%; Score 233; DB 1; Length 478;
 Best Local Similarity 23.7%; Pred. No. 1e-07;
 Matches 109; Conservative 92; Mismatches 174; Indels 84; Gaps 20;
 19 IGYGIATFVGLGMAFAPLDSAVIANGVVS-AEVSODVOHLEGGMLAKTIVEGKV 77
 Db 60 IAY-LIMFLVVAIVLASKEYEIVATAGKLTGSGRSKEIPIMALVQELFVVDGQFV 118
 QY 78 KAGVLFELDPTQANAAGITRNQ-YVALKAME---ARLLA-BRDQPSISPPADLT- 129
 Db 119 EKGQLVSL--TALGSDADIKKTMASTSLAKIENRYOTLLTAIKESLPVT---DLR 172
 QY 130 -----SQRADPVVARAIADDAQFTERQTIQGVDDLMNAOR-----LQYSEIEGID 177
 Db 173 TEFKDSEEDRLRIKHLIEQYTTWQKQT---QKTLVKKREAKQTTFAVRYREGAT 229
 QY 178 ROTGGLKQGLFIDELELDLRLKLDKGLVPRRLALEARAGLSGSGITRLADRSKAVQ 237
 Db 230 RLEQ-----EKPRFKALVYKQKSKSKHELLAQE---NKLLEAQNELAVYSKLNE 276
 QY 238 GASD-----TQLKVRQIKOEFEQVQSQSI-TETRVRLAEVTEKEVVASDAOKRIKIV- 288
 Db 277 LENDLLNVKEELLEITQFESDVLKELKQHTENRQLRL-----ELEKNNQRQASMR 330
 QY 289 SPVNGTAQNLREFTEGAVVRAEPLVDIAPDEAFVIOAHFOPTVDNVHMGVTEVRLP 348
 Db 331 APVSGTVQQLKHTITGAVVTAETLMIITVPEEDVLEATLVPNKDIGFAVAQOEVIKVE 390
 QY 349 AF-----HSAG-----NPDEPRHDPAVADRISDPQKARLFLGIVRVVVKQPLRL 395
 Db 391 TFPYTRYGYLGRKIKHISPDALQEPNGL-----VFNAITAIADKRLNLSPOGR 438
 QY 396 RGRVTAGMPAQVIPTGERTVLOYLFSPRLDRLRTTMR 434
 Db 439 KIDLSGMSVTAETIKTGRSVISYLLSPLEESVSESLRE 477


```

Db 44 VSGEVTTPRAVNIYSGVGFVRFVHEGOLIKKDPYLLIDSKS-----TRN---- 93
OY 105 LKAMEARLLAERDORPSISFPADLTSQ--RADPMVARATADEQAQFTEEROTIGOVADM 162
Db 94 -----GIYTDNHR-----DIENQLVKVNITSR--LEESKIT-----LDTL 129
OY 163 NNAORLY-----OSEIEGIDROFOGKIDOLFIEDELIDKLKLYDKGLVPPRL---LA 213
Db 130 EKQRRLQYIDAFRRSSDI--IQRAEGIK-----IMKNMNYVYOSKGLINKDQLNVOYA 183
OY 214 L-----EARGSLSG-----SIGRLTADRKAQVAGASDTOLKVRQIKOEFEVOSITETR 265
Db 184 IYQQQNNLISGQNEQNALQITLLESOIQTAADFNRITQYOMELQRL----- 233
OY 266 VALAEVEKEVVAQAQRIRKIVSPVNGAOLRFTTEBAVYRAAEPLVDIAPDEAFYI 325
Db 234 -----LQKELVNTVEGEIIRALSDGKVDLS--VTQGMVNTGDSLQIVPENIENY 286
OY 326 QAHPQPTD-VDNVHGMVTEVRLPAFHS-----AGNPPD 358
Db 287 LILWMPNDAVPIYISAGDKVNIRYEAFPSSEKFGFSATVKTISRPASTQEMLTYGAPQN 346
OY 339 EKHDPVNAADRISDPQKARLFLGIVRDVQOLPPLHGRVTAGAPQVIVPTGERTVIO 418
Db 347 TEGASVPMYKVIATPEKQ-----IIRYDEKYL-----LENGMAESTLFLKRRRIYO 394
OY 419 YLESPLRD 426
Db 395 WMLSPFYD 402

```

RESULT 13

YXCR_ECOLI

ID YXCR_ECOLI STANDARD: PRT: 343 AA.

```

AC P32716;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein yjcr.
GN yjcr OR B4082.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=94089392; PubMed=8263357;
RA Blattner F.R., Burland V.D., Plunkett G. III, Sofia H.J.,
RA Daniels D.L.;
RT "Analysis of the Escherichia coli genome. IV. DNA sequence of the
RT region from 89.2 to 92.8 minutes."
RL Nucleic Acids Res. 21:5408-5417(1993).
CC -1- SUBCELLULAR LOCATION: Inner membrane-bound (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE HLYD FAMILY OF SECRETION PROTEINS.
CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U00006: AAC43176.1;
DR EMBL: AE000481: AAD13465.1;
DR EcoGene: EG1954; yjcr.
DR InterPro: IPR002215; HLYD.
DR Pfam: PF00529; HLYD.
DR TrEMBL: TIGR00998; Bae0101; 1.
KW Hypothetical protein; Transmembrane; Inner membrane; Transport;
KW Complete proteome.
FT DOMAIN 1 12 CYTOPLASMIC (POTENTIAL).

```

```

FT TRANSMEM 13 33 POTENTIAL.
FT DOMAIN 34 343 PERIPLASMIC (POTENTIAL).
SQ SEQUENCE 343 AA: 36908 MW: C59988707C475FFE CXC64.

```

Query Match: 6.8%; Score 147.5; DB 1; Length 343;
 Best Local Similarity 22.5%; Pred. No. 0.016;
 Matches 78; Conservative 53; Mismatches 131; Indels 85; Gaps 14;

```

OY 4 PRTQPTDNFQAVARIGVGIILTFVGLGMAFAFPLDSAVANGVSAVESQVQHLEG 63
Db 5 PK-KAPRSKFPALLVALVALVALVTV--TW--VDSAPSTNAYVASADTIDVAVEVS 56
OY 64 GIALKILVREGKAVAGVLFELDPT--QANAAGITRQVYALKAMEARLLAERDQPS 121
Db 57 GRIVELAVTDNVAQGGDLFRIDPRPYANLA-----KAESLAIDQIMLTQRSVA 111
OY 122 ISFPADLTSQADPMVARAIADEQAQFTEEROTIGOVADMNAORLYOSEIEGIDROTO 181
Db 112 QQFGAD--SVNATVEKARAA--KQATDTLRR-----TE 141
OY 182 GLKDQGFIEDELIDRLKLYDKGLVPRRLALLEAR--AGSLSGSIGRLTADRKAQVGS 240
Db 142 PLTKE-GEVSADVDPAKTAQRAEADLNAVLQASASAVSGDALVAQRA----- 193
OY 241 DTQLKVRQIKOEFEVOSITETRVLAETVEKEVVASDAQRIKIYSPVNGTAQNLRF 300
Db 194 -----AVENDIALTKHLEMAI-----VRAPFDGRVYISLK- 223
OY 301 FTGAVVRAAEPLVDIAPDEAFVIOAHFQPPDVNDVHGMVTEVRL 347
Db 224 TSVGQFASAMRPIFTLIDTRHWYVI-ANRETDLNIRSGTPATIRL 269

```

RESULT 14

EXOF_RHIME

ID EXOF_RHIME STANDARD: PRT: 421 AA.

```

AC Q02728;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Exopolysaccharide production protein exoC precursor.
GN EXOF OR RB1068 OR SMB20945.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RCR2011 / SU47;
RX MEDLINE=93177026; PubMed=8439670;
RA Mueller P., Keller M., Weng W.M., Quandt J., Arnold W., Puehler A.;
RT "Genetic analysis of the Rhizobium meliloti exoC operon: ExoC is
RT homologous to sugar transferases and ExoD represents a transmembrane
RT protein."
RL Mol. Plant Microbe Interact. 6:55-65(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396508; PubMed=11481431;
RA Finan T.M., Weidner S., Wong K., Buhmester J., Chain P.,
RA Vorholter F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
RA Golding B., Puehler A.;
RT "The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-
RT fixing endosymbiont Sinorhizobium meliloti."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
CC -1- FUNCTION: INVOLVED IN SUCCINOGLYCAN (EPS I) SYNTHESIS. NEEDED FOR
CC THE ADDITION OF THE RING SUGAR (GALACTOSE) TO THE ISOPRENOLID
CC CARRIER.
CC -1- PATHWAY: Exopolysaccharide biosynthesis.
CC -1- SUBCELLULAR LOCATION: Periplasmic (probable).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```


CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L05588; AAA26265.1; -
 DR EMBL: AL003645; CAC949468.1; -
 KW Exopolysaccharide synthetase; Plasmid; Periplasmic; Signal;
 FT Complete proteome.
 FT SIGNAL 1 31 POTENTIAL.
 FT CHAIN 32 421 EXOPOLYSACCHARIDE PRODUCTION PROTEIN
 FT EXOF.
 SQ SEQUENCE 421 AA; 45871 MM; 55D8EC55B3A469EF CMC64;
 Query Match 6.7%; Score 146; DB 1; Length 421;
 Best Local Similarity 22.3%; Pred. No. 0.026;
 Matches 66; Conservative 47; Mismatches 121; Indels 62; Gaps 6;
 QY 22 GILALTFGLGMAFAFLDSAVINGVSAVESODVHLEGGMLAKI----- 69
 Db 79 GSLSLPEFGDLP-----ASGKTSEVAEEI---GVKKQKLGKDRPSASVE 122
 QY 70 -----LVREGKVKAGQ-----VL-----FELDPQANNAAGIT 98
 Db 123 MAQRYPVYLSEVOTPGETPYAPNLTLYKANVSLGGGLRADNGOFARDYINASESAVQ 182
 QY 99 RNQYVALKAMEARLLAEERDQRPISFPADLTSORADPNVARIADDEQAQFERRQTOGO 158
 Db 183 VAERSRLIRARRLLAEIGKRDITPMPELKNV---PDAEKILDEBTALMESRQKQK 239
 QY 159 VDLNMAQQLQYQSEIEGIDRQTOGKLDGFIEDLIDRLKLYDKGLVPRRLALEKRA 218
 Db 240 LDALADLFSLLQSELEAKAKAEQARQLATELRDRVDLSAEKGLALSQRKLSLEQRY 299
 QY 219 GSLSGISRLTADRSKAVOGASDTQKVKROIKEFEFQVOSITETRYRLAEVTEK 274
 Db 300 ADVQASLIDITATSLAKAKQDASKAQAQDETNRNDMDQALQELQNTAEELDTITLK 355
 RESULT 15
 YHII_ECOLI STANDARD; PRT; 355 AA.
 AC P37626;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein yhiI precursor.
 GN YHII OR B3487.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE-94316500; Pubmed-8041620;
 RA Sofia H.U., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;
 RT "Analysis of the Escherichia coli genome. V. DNA sequence of the
 RT region from 76.0 to 81.5 minutes."
 RL Nucleic Acids Res. 22:2576-2586(1994).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U00039; AAB18463.1; -
 DR EMBL: AF000424; AAC76512.1; -

DR EcoGene: EG12224; yhiI.
 DR InterPro: IPR002215; HlyD.
 DR Pfam: PF00529; HlyD; 1.
 KW Hypothetical protein; Signal; Complete proteome.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 355 HYPOTHETICAL PROTEIN YHII.
 FT EXOF.
 SQ SEQUENCE 355 AA; 38828 MM; 6784DF1F19D5FPA CMC64;
 Query Match 6.6%; Score 142; DB 1; Length 355;
 Best Local Similarity 21.2%; Pred. No. 0.037;
 Matches 87; Conservative 65; Mismatches 158; Indels 100; Gaps 20;
 QY 22 GILALTFGLGMAFAFLDSAVINGVSAVESODVHLEGGMLAKITVREGKVK 78
 Db 14 GILAV--AAIYAWMLLRAPGYPEGFAVNGKIEA-TEVDINASKIAGRITDTLVKEGKVR 70
 QY 79 AGOVLFEIDPQANNAAGITRNQYVALKAMEARLLAEERDQRPISFPADLTSORADPNVA 138
 Db 71 EGEVLAKMD-----TR-----VLQEQVLEMAIQKKEAQ----- 99
 QY 139 RAIDDEQAQFERRQTOGQVADLNNARLQYQSEIEGI-DKQTOGLK-DQLGFIEDLID 196
 Db 100 -AVAAQAQLLEORQSETRAAQSLVN---QRAEILDSYAKKRRHTRSRSLAQGAISAQQLD 154
 QY 197 LRKLVDKGLVPRRLALEARAGSLSGISGRITFADRSKAVOGASDTQKVKROIKEFEQ 256
 Db 155 -----DRAAASARAALESKAAQVSAKAAIEARNTIIQ----- 190
 QY 257 VSQITETRYRLAEVTEKEVVAASDAQRIKIVSPVNGTAQMLRFELEGAVYRAAELVDI 316
 Db 191 -----AQTRVBAQAQATERIAADIDDSLEK--APRDSGVQ-YRAAPEVGLAAGRVLM 242
 QY 317 APEDEAVYQAHFOPTD-VDNVHMG-----MTEVRLPA---FHSAGNPDEKHP 363
 Db 243 V--DLSVYMTFFLPTEQAGTLKGGEARLILDAAPDIRIPATISFVSAQFTPK---T 297
 QY 364 VAVADRISDPKQARLFLGIYRVYKOLPRLRGVYTGMAQYIVPTGE 413
 Db 298 VETSD-----ERLKMFRVKARIPPELLQHL-EYKTLDPGVAMVRVNE 341

Search completed: April 22, 2003, 12:31:38
 Job time : 28 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:29:17 ; Search time 21 Seconds
(without alignments)
1991.356 Million cell updates/sec

Title: US-09-913-414-5

Perfect score: 2165

Sequence: 1 MKPPKIQRPDMFQAVARIG.....VLQYLFSPRLRDTLTFTMREE 435

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

1: PIR_73:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1984.5	91.7	436	2	B87374
2	589	27.2	474	2	B98370
3	584	27.0	437	2	A82972
4	583	26.9	439	2	C96003
5	568	26.2	434	2	A83099
6	568	26.2	436	2	G98187
7	564	26.1	443	2	B83221
8	543.5	25.1	473	2	C95953
9	542.5	25.1	437	2	H97683
10	542.5	25.1	437	2	A12908
11	530.5	24.0	511	2	D97753
12	519.5	24.0	511	2	D71687
13	514.5	23.8	432	2	G26697
14	514.5	23.8	432	2	G83489
15	506.5	23.4	448	2	S12526
16	476.5	22.0	442	2	A50477
17	471.5	21.8	437	2	B49933
18	438.5	20.3	452	2	A82965
19	438.5	20.3	497	2	G98318
20	368.5	17.0	387	2	A60835
21	368	17.0	481	2	D82381
22	367	17.0	395	2	D83412
23	329	15.2	475	2	B81047
24	325	15.0	391	2	D85547
25	325	15.0	391	2	H90696
26	314	14.5	473	2	P82561
27	308.5	14.2	478	1	S48045
28	288.5	13.3	478	1	LEBCD
29	288	13.3	507	2	AD2140

30	267.5	12.4	477	2	B61378
31	267.5	12.4	478	2	S10058
32	266	12.3	440	1	BYBRCD
33	264	12.2	438	2	T03513
34	262	12.1	478	2	E43599
35	261.5	12.1	454	2	G81828
36	260	12.0	500	2	A82517
37	251.5	11.6	479	2	T00229
38	251	11.6	467	2	G82198
39	248	11.5	498	2	AH2468
40	234	10.8	478	2	D30169
41	215.5	10.0	583	2	S75805
42	203.5	9.4	425	2	AB1018
43	201.5	9.3	425	2	T14873
44	196	9.1	455	2	D95923
45	194	9.0	512	2	AB2047

ALIGNMENTS

RESULT 1

B87374

RsaA secretion system, membrane protein RsaA [Imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C:Accession: B87374

R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwin, M.L.; Haft, D.H.; Ko

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete genome sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: B87374

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-436 <STO>

A:Cross-references: GB:AE005673; NID:q13422299; PIDN:AKK2293.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC1009

Query Match	91.7%	Score 1984.5;	DB 2;	Length 436;
Best Local Similarity	93.6%	Pred. No. 7.9e-116;		
Matches 408;	Conservative 3;	Mismatches 24;	Indels 1;	Gaps 1;
QY	1			
DB	1			
QY	60			
DB	61			
QY	120			
DB	121			
QY	180			
DB	181			
QY	240			
DB	241			
QY	300			
DB	301			
QY	360			
DB	361			

OY 420 LFSPLRLRTTMRRE 435
 Db 421 LFSPLRLRTTMRRE 436

RESULT 2

F98310

rhizobium secretion protein rspe (AF141932) [imported] - Agrobacterium tumefaciens (st
 C:Species: Agrobacterium tumefaciens
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
 C:Accession: F98310

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quirillo, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
 Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tun
 A:Reference number: A97359; PMID:11743194
 A:Accession: F98310

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-474 <KUR>
 A:Cross-references: GB:AE007870; PIDN:AAK90008.1; PID:915159979; GSPDB:GN00170
 C:Genetics:
 A:Gene: AGR_L_2881
 A:Map position: linear chromosome

Query Match 27.2%; Score 589; DB 2; Length 474;
 Best Local Similarity 31.9%; Pred. No. 3.2e-29;
 Matches 140; Conservative 85; Mismatches 210; Indels 4; Gaps 2;

OY 1 MKPPIQPTNFOAVAR---IGYGIALTFFVGLGNAAFPLDSAVYANG-VYSAEVSQ 56
 Db 36 MRMRRIDORTVYSRSIRKHLIAGLAGAALLAGVGMAATTNLGAVVASGHLVDSYTK 95
 OY 57 DVQHLEGLAKIIVREGEKYGAGVLFELDPQTQANAAGITRNQYVALKMEARLLAER 116
 Db 96 KVGHPKGVGEILLVNEGERKAGDVWMRLDATQTRANLAVTRKRLDELGARMARLEAER 155
 OY 117 DQRSISPPADLTSGRADPMYARAIADBOAQFTERRQTIQGVDMNQRLOYSEIEIGI 176
 Db 156 DDLAEIIPQALIRRDPPDVASAMRSETKLFEEFKSYREGRKQALERTIQFHEIEIGL 215
 OY 177 DRQIOLKDOGFIEDELIDRLKLYDKGLVPRRLALEARAGSLSSIGLTDREKAV 236
 Db 216 KAQEVADNGLAIVLEAETTSOKSLREGVSVORLNSLQTOAATFGGEREKIAVQAOA 275
 OY 237 QGASDTOLKVRQIKOEFEEQVSQSTITERVLAETVEKEVVASDAOKRIKIVSPVNGTAQ 296
 Db 276 GRITETKQIQLIDQELTEVGRLEIQAGMGEVERKVAEDELKRIIDIVAQSGMVH 335
 OY 297 NLRFETGSAVRAAPLVYDIAPDEDAFYQAHFOPTDVNVHMGVTEVRLPAHSAGNP 356
 Db 336 EMAVHTVGVVTPADPILILYPPDGDLEALVQIVPKIDQLOVQOKAMLRMTAFNQRTVP 395
 OY 357 DPERHDPVAVADRISSDPQKARLFGLIVRVYKOLPRHLRGRVNAQPAQVIYVPGERTV 416
 Db 396 ELEGVASIAADITTDQRTGSLYLARISVPSEKELNNAPLVGPMAEAFIQTSEKTA 455
 OY 417 LQYLFSPRLRLRTTMRRE 435
 Db 456 LSYIAKPLTDQISRAREE 474

RESULT 3

AE2972

secretion protein, HlyD family [imported] - Agrobacterium tumefaciens (strain C58, Dupont
 C:Species: Agrobacterium tumefaciens
 C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002

C:Accession: AE2972
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McCell
 ; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A:Reference number: AB2577; PMID:11743193

A:Accession: AE2972
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-437 <KUR>
 A:Cross-references: GB:AE008689; PIDN:AAU44195.1; PID:917741773; GSPDB:GN00187
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: prse
 A:Map position: linear chromosome

Query Match 27.0%; Score 584; DB 2; Length 437;
 Best Local Similarity 31.9%; Pred. No. 5.8e-29;
 Matches 139; Conservative 84; Mismatches 209; Indels 4; Gaps 2;

OY 4 PKIQPTNFOAVAR---IGYGIALTFFVGLGNAAFPLDSAVYANG-VYSAEVSQ 59
 Db 2 PNIDORTVYSRSIRKHLIAGLAGAALLAGVGMAATTNLGAVVASGHLVDSYTKVQ 61
 OY 60 HLEGGLAKIIVREGEKYGAGVLFELDPQTQANAAGITRNQYVALKMEARLLAERDOR 119
 Db 62 HPKGVGEILLVNEGERKAGDVWMRLDATQTRANLAVTRKRLDELGARMARLEAERDL 121
 OY 120 PSISFPADLTSGRADPMYARAIADBOAQFTERRQTIQGVDMNQRLOYSEIEIGIDRQ 179
 Db 122 AELIIPQALIRRDPPDVASAMRSETKLFEEFKSYREGRKQALERTIQFHEIEIGL 181
 OY 180 TQGLKDOGFIEDELIDRLKLYDKGLVPRRLALEARAGSLSSIGLTDREKAVQA 239
 Db 182 EVAYDNGLAIVLEAETTSOKSLREGVSVORLNSLQTOAATFGGEREKIAVQAOAGRI 241
 OY 240 SDTOLKVRQIKOEFEEQVSQSTITERVLAETVEKEVVASDAOKRIKIVSPVNGTAQNL 299
 Db 242 TETKQIQLIDQELTEVGRLEIQAGMGEVERKVAEDELKRIIDIVAQSGMVHMA 301
 OY 300 FTEGAVVRAAPLVYDIAPDEDAFYQAHFOPTDVNVHMGVTEVRLPAHSAGNDPE 359
 Db 302 VHTGVGVVTPADPILILYPPDGDLEALVQIVPKIDQLOVQOKAMLRMTAFNQRTVELE 361
 OY 360 RHDVAVADRISSDPQKARLFGLIVRVYKOLPRHLRGRVNAQPAQVIYVPGERTV 419
 Db 362 GHSRIADITTDQRTGSLYLARISVPSEKELNNAPLVGPMAEAFIQTSEKTA 421
 OY 420 LFSPLRLRTTMRRE 435
 Db 422 IAKPLTDQISRAREE 437

RESULT 4

C96003

protein secretion protein, HlyD family [imported] - Sinorhizobium meliloti (strain 10
 C:Species: Sinorhizobium meliloti
 C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001

C:Accession: C96003
 R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
 A:Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing e
 A:Reference number: A95842; MUID:21396508; PMID:11481431
 A:Accession: C96003

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-439 <KUR>
 A:Cross-references: GB:AL591985; PIDN:CAC49691.1; PID:915141178; GSPDB:GN00167
 A:Experimental source: strain 1021, megaplasmid PSYMB
 R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Batloy-Hudl
 pella, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.
 L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau
 behault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,
 A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A:Accession: H97683
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-437 <KUR>
 A:Cross-references: GB:AE007869; PIDN:AAK88425.1; PID:q15157919; GSPDB:GN00169
 C:Genetics:
 A:Gene: AGR_C_4904
 A:Map position: circular chromosome

Query Match 25.1%; Score 542.5; DB 2; Length 437;
 Best Local Similarity 30.8%; Pred. No. 2.2e-26;
 Matches 130; Conservative 83; Mismatches 204; Indels 5; Gaps 2;

15 AAVRIGYGIATLTVGLGMAAFAPLDSAVIANGVSAS-VSODVOHLEGMLAKITVRE 73
 DB AVVAVIGLGLL---VGMGMAFAPKLAGAVATGRVVEGNSKKIHLGSGIYSEINVE 75
 74 GKKVAGOVLFELDPQANAAGITRNOYVALKAMEARLLAERDORPSISFPADLTQRA 133
 DB GGRVAGGQILRLSATVVOANLSTIENTLAOLYSRRARAEIAEERPSFTVEDLTALTS 135
 134 DPMVRAIADDEQAOFERROTIOGOVDLMNAORLOYQSEIEGIDROTOGLKOLGFIENE 193
 DB SSKATFIDSEONLNSRNALIGMKQLATRLQADARGLDVOVETEMELATVKED 195
 194 LIDRLKLYDKGLVPRRLALAEARAGSLSGISGRLTADRSKAVOGASDTOLKVRQIKOFE 253
 DB VSKTDELLKKGIVTLQRLNLKRLNSLEGQGOYIAAQAQVYVSKLSELDQLQDEDR 255
 254 FEQVQSITETRVRLAEVTEKEVVASDAOKRIKIVSPVNGTQONLRFEEGAVVRAEPL 313
 DB KSEVTKDLTISBATVAEYERLAATRDQDLRIDIRSPIAGRITOLSVHNINGVIOGFEVL 315
 314 VDIAPDEAFVIOAHFOPTVDNVHMGWTEYRLPAFHSAGNDPDRHDPVAVADRISDP 373
 DB MLVPEKDDLALEANTTPDDIOIYVGPYVFTAFNGSTTPDLSAEVAVAPDLQOTS 375
 374 OKQARLFLGIVRVYKQPLRLGRVTAQMPAOVIVPTGERTVLOYLFSPLDITLTNR 433
 DB RGTSTYVLRIRPNKMGHLPGLKLYPGMPAEVFIQTSESVLSYFVAPFDRLKKTIV 435
 434 EE 435
 DB 436 QE 437

RESULT 10

AI2908
 HlyD family secretion protein [imported] - Agrobacterium tumefaciens (strain C58, Dupont
 C:Species: Agrobacterium tumefaciens
 C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
 C:Accession: AI2908
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, H.
 et al.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan,
 Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 ster, E.W.
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A:Reference number: AB2577; PMID:11743193
 A:Accession: AI2908
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-437 <KUR>
 A:Cross-references: GB:AE008688; PIDN:AAL43687.1; PID:q17741213; GSPDB:GN00186
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: prsE
 A:Map position: circular chromosome

Query Match 25.1%; Score 542.5; DB 2; Length 437;
 Best Local Similarity 30.8%; Pred. No. 2.2e-26;
 Matches 130; Conservative 83; Mismatches 204; Indels 5; Gaps 2;

15 AAVRIGYGIATLTVGLGMAAFAPLDSAVIANGVSAS-VSODVOHLEGMLAKITVRE 73
 DB AVVAVIGLGLL---VGMGMAFAPKLAGAVATGRVVEGNSKKIHLGSGIYSEINVE 75
 74 GKKVAGOVLFELDPQANAAGITRNOYVALKAMEARLLAERDORPSISFPADLTQRA 133
 DB GGRVAGGQILRLSATVVOANLSTIENTLAOLYSRRARAEIAEERPSFTVEDLTALTS 135
 134 DPMVRAIADDEQAOFERROTIOGOVDLMNAORLOYQSEIEGIDROTOGLKOLGFIENE 193
 DB SSKATFIDSEONLNSRNALIGMKQLATRLQADARGLDVOVETEMELATVKED 195
 194 LIDRLKLYDKGLVPRRLALAEARAGSLSGISGRLTADRSKAVOGASDTOLKVRQIKOFE 253
 DB VSKTDELLKKGIVTLQRLNLKRLNSLEGQGOYIAAQAQVYVSKLSELDQLQDEDR 255
 254 FEQVQSITETRVRLAEVTEKEVVASDAOKRIKIVSPVNGTQONLRFEEGAVVRAEPL 313
 DB KSEVTKDLTISBATVAEYERLAATRDQDLRIDIRSPIAGRITOLSVHNINGVIOGFEVL 315
 314 VDIAPDEAFVIOAHFOPTVDNVHMGWTEYRLPAFHSAGNDPDRHDPVAVADRISDP 373
 DB MLVPEKDDLALEANTTPDDIOIYVGPYVFTAFNGSTTPDLSAEVAVAPDLQOTS 375
 374 OKQARLFLGIVRVYKQPLRLGRVTAQMPAOVIVPTGERTVLOYLFSPLDITLTNR 433
 DB RGTSTYVLRIRPNKMGHLPGLKLYPGMPAEVFIQTSESVLSYFVAPFDRLKKTIV 435
 434 EE 435
 DB 436 QE 437

RESULT 11

D97753
 alkaline protease secretion protein AprE [imported] - Rickettsia conorii (strain Mali
 C:Species: Rickettsia conorii
 C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 30-Sep-2001
 C:Accession: D97753
 R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.;
 Science 293, 2093-2098, 2001
 A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
 A:Reference number: A97700; MUID:21442074; PMID:11557893
 A:Accession: D97753
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-511 <KUR>
 A:Cross-references: GB:AE006914; PIDN:AAL02966.1; PID:q15619497; GSPDB:GN00173
 C:Genetics:
 A:Gene: aprE

Query Match 24.5%; Score 530.5; DB 2; Length 511;
 Best Local Similarity 30.0%; Pred. No. 1.5e-25;
 Matches 135; Conservative 88; Mismatches 202; Indels 25; Gaps 7;

5 KIORPDNQAAV--IGYIATLTVGLG--MAAFAPLDSAVIANGVSAS-ODVO 59
 DB KIDKDRNNVAAARSPILFGLYVIFVLIGLMSALAPLDSGAVAVGYMSTNKKTIQ 127
 60 HLEGMLAKITVREGKRVAGOVLELDPQANAAGITRNOYVALKAMEARLLAERDOR 119
 DB HHEGGIIMAYVQGGKVEGDKLIELETRIKSEHENILGGRNPLATENRIABRONL 187
 120 PSTSPADLTQRAQPMVRAIADDEQAOTERROTIOGOVDLMNAORLOYQSEIEIDQ 179
 DB EQIEFSDFLMODINPEVAKIITHQENLRSRKEVYNSKDLHONIAOLEKKEIELEAK 247
 180 TQGLKQOLGFIEDDELKLYDKGLVPRRLALAEARAGSLSGISGRLTADRSKAVQA 239
 DB KYAASKTAVYODRLKALFTLKEKGVOKAALLDDEAKAASKSDVATTEAELAGRAHI 307
 240 SDTOLKVRQIKOFEFEQVQSITETRVRLAEVTEKEVVASDAOKRIKIVSPVNGTAQNLR 299

Db 308 TETQIKITNOONKYTEERITLTELREAOVOTASIKERKNALTSJLKNVILIRAPDGVNSIK 367
 QY 300 FFTGAIVRAAEPLVDIAPEDEAFVIOAHFOPTVDVNMGMVTEVRLPAFHSAGNP--- 356
 Db 368 YHTIGVISHQOPIMEISPTNDPLILEKAVSOKNIDSHVGLVAKIRSAKSRTPPTFT 427
 QY 357 -DPERHDPVAVADRISDF-OKOARLEGLIVND-----VKQLPRLRGVTAQMPA 405
 Db 428 GKVVISISPDIVQDERQYQGGQODNYYVARVEIDMEFNKAVAKNLELH-----PGMOA 481
 QY 406 QVIVPTGERTLOYLFSPRLRTKTRREE 435
 Db 482 EVQIVTGTIRLLRLDDEVTDPAFAKREK 511

RESULT 12

D1687
 alkaline proteinase secretion protein apre (apre) RP314 - Rickettsia prowazekii
 C:Species: Rickettsia prowazekii
 C>Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000
 C:Accession: D1687
 R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Slicheritz-Ponten, T.; Alsmark, U
 Nature 396, 133-140, 1998
 A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
 A:Reference number: A71630; MUID:99039499; PMID:9823893
 A:Accession: D1687
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-511 <AND>
 A:Cross-references: GB:A235271; GB:A235269; NID:93868717; PIDN:CAAL4774.1; PID:9386087
 A:Experimental source: strain Madrid E
 C:Genetics:
 A:Gene: apre; RP314

Query Match 24.0%; Score 519.5; DB 2; Length 511;
 Best local similarity 29.3%; Pred. No. 7.1e-25;
 Matches 134; Conservative 93; Mismatches 192; Indels 39; Gaps 10;

QY 6 IQRPD---NFQAVAR--IGYIATLFGVLG--MAAFAPLDSAVANGVSAEVS-Q 56
 Db 65 ITRKTDEMRNVAAQARSPLFGIYVIFLVIGLMSALAPLDSGAVAGIVIPSTKK 124
 QY 57 DVQHLGGMLAKIIVREBEKRAQVLELDPTQANAAAGITRQYVALKMEARLAE 116
 Db 125 TIQNEGGIINAIYVQGGDKVEGDKLELETRKSEHENTLOQRNFLTENKRLAE 184
 QY 117 DQRSISFPADLTQSRAD-PWVARAIADQEQAFERRQTIQGVYDLMAQRLQYQSEIEG 175
 Db 185 DNLQIIF-SDFLMQNTINLPVAKIITHQENLFKSRKVEYSEKDAFNQNTAQLEKIDG 243
 QY 176 IDROTQGLKQDGFIEDELIDLRKLYDKGLVPRRLALLEARAGSLSGISRLTADRSKA 235
 Db 244 LEAKRIAAKSTSEYVQDRKALRFLKEGFIQKALLQEQAKVAASDVATTEAETAGI 303
 QY 236 VQGSDDPOLKVRQKQEFEEVQSITETRVRLAEVTEKEVAVADAKRIKIVSPVNTA 295
 Db 304 RHAITETQIKINHONKYTEERITLTELRAQIOTASLKEKYNSTDLNRIIRSPVDIV 363
 QY 296 QNLRFEGAVVRAAEPLVDIAPEDEAFVIOAHFOPTVDVNMGMVTEVRLPAFHSAGN 355
 Db 364 NNLKYYHIGVISHQOPIMEISPTNDPLILEARISOKNIDSHVGLVAKIRSAKSRTPFT 423
 QY 356 P-----DPERHDPVAVADRISDPQKARLEGLIVND-----VKQLPRLRG 397
 Db 424 PTFGKVVISPD---IVDERQHLQGGQODNYYVARVEIDMEFNKAVAKNLELH--- 476
 QY 398 RVTGMPAQVIVPTGERTLOYLFSPRLRTKTRREE 435
 Db 477 ---PGMAEVOIVTGTIRLLRLDDEVTDPAFAKREK 511

RESULT 13

S26697
 apre protein - Pseudomonas aeruginosa
 C:Species: Pseudomonas aeruginosa
 C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 31-Mar-2000
 C:Accession: S26697
 R:Duong, F.; Lazdunski, A.; Cami, B.; Murgier, M.
 Gene 121, 47-54, 1992
 A:Title: Sequence of a cluster of genes controlling synthesis and secretion of alkali
 A:Reference number: S26696; MUID:93051361; PMID:1427098
 A:Accession: S26697
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-432 <DQ>
 A:Cross-references: EMBL:X64558; NID:945279; PIDN:CA45856.1; PID:945281
 C:Superfamily: hemolysin secretion protein D; lipoyl/lipotin-binding homology
 C:Keywords: transmembrane protein
 F:48-91,284-313/Domain: lipoyl/lipotin-binding homology #status atypical <LPB>

Query Match 23.8%; Score 514.5; DB 2; Length 432;
 Best local similarity 31.6%; Pred. No. 1.2e-24;
 Matches 137; Conservative 88; Mismatches 182; Indels 27; Gaps 11;

QY 15 AVARIGVGIATLTFVGLGMAAPALDS--AVANGVSAEVSQDVOHLEGMLAKILVR 72
 Db 11 AYARGLWLVLFEGGALMAFAFAPLDGVAVPATVITISGO-KSVQHPDLGGVAKHILVR 69
 QY 73 EGEKKAQGVLELDPTQANAAAGITRQYVALKMEARLAEERDQPSISFPADLTQSR 132
 Db 70 DQGVHVEGEPLIRHEPQARAVNSDLNRYANARHLNARLQAEVDGRTLEMPRLAEQA 129
 QY 133 ADPVAVARIADQEQAFERRQTIQGVYDLMAQRLQYQSEIEGIDROTQ---LKQDLG 188
 Db 130 PLPTLGERL-ELQROLHRSROTALANELSLRANIEGLRAQLEGI-ROTEGNORLQRL- 186
 QY 189 FIEDELIDLRKLYDKGPRRLALE-----ARAGSLGSGRLTADRSKAVQASD 241
 Db 187 -LNSQLSARLAEEGYPRNQLLEQEQRLAEVNAKRISSESGRGQIR-----QSIAE 238
 QY 242 TQLKVRQIKQEFEEVQSITETRVRLAEVTEKEVAVADAKRIKIVSPVNGTQNTFF 301
 Db 239 AQMRIAQREERYREVNQALQETQVNNARTLWEELSSAREYELRHAIRPAGVYAGLKF 298
 QY 302 TEGAVVRAAEPLVDIAPEDEAFVIOAHFOPTVDVNMGMVTEVRLPAFHSAGNPDERH 361
 Db 299 TDGGEVIGGELIMYVPSDSLEVEGLAVNLVRIHSGLEVEMLFTAFQNSKPRVATGE 358
 QY 362 DPVAVADRISDPQKARLEGLIVRVYQKLPRLRG-RVTGMPAQVIVPTGERTVLOYL 420
 Db 359 VTMVSADRLDEQKKQPYRVRAQVDAAM-CKLKGIDIRPMNAVQVIVRTGERSLNTYL 417
 QY 421 FSPRLDRLTTRRE 434
 Db 418 FKPLFDRHVALAE 431

RESULT 14

G83489
 alkaline proteinase secretion protein Apre PA1247 [imported] - Pseudomonas aeruginosa
 C:Species: Pseudomonas aeruginosa
 C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: G83489
 R:Stover, C.K.; Pham, X.Q.; Eyrin, A.L.; Micoqueli, S.D.; Warren, P.; Hickey, M.J.;
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; L
 .; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
 A:Reference number: A82950; MUID:20437337; PMID:10984043
 A:Accession: G83489
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-432 <STO>
 A:Cross-references: GB:AE004554; GB:AE004091; NID:99947174; PIDN:AAQ04636.1; GSPDB:GN
 A:Experimental source: strain PA01

C:Genetics:
A:Gene: aprE; PA1247

Query Match 23.8%; Score 514.5; DB 2; Length 432;
Best Local Similarity 31.6%; Pred. No. 1.2e-24;
Matches 137; Conservative 88; Mismatches 182; Indels 27; Gaps 11;

15 AVARIGYIALTFVGLGMAFAPLDS--AVIANGVSAEVSQDVQHEGGLAKILVR 72
11 AVARIGYIALTFVGLGMAFAPLDS--AVIANGVSAEVSQDVQHEGGLAKILVR 69
73 EGEKKAQGVLELPTQANAAAGITRNOYVALKAMEARLARERDQPSISFPADITSOR 132
70 DGOHVEAEPLTRMEPTQARAVNDSLNRNANARLQAEYDGRRTLEMPAGIAEQA 129
133 ADPMVARIADDEQAQFTERROT-IOGVDMNAQRLQOSEIEGIDRQOG---LKDOLG 188
130 PLPTIGERL-ELQROLHSROTALNELSALNANIEGLRAQLEGL-RQTEGNORLOQRL- 186
189 FIEDELIDRLKLYDKGLVPRRLALE-----ARAGSLSGSIGRLTADRKAQVAGASD 241
187 -LNSQLSGARDLAEGVYPRNOLLEOERQLAENVNARLSSSSRFQOIR-----QSTAE 238
242 TOLKVRQIKQEFFEQVQSITETRVRLAEVTEKEVVAQAQRIRKIVSPNGCTAQNLEEF 301
239 AQMRIQREEEYRKEVENCQLAETQVNAFTLMEELSSARVELRHAETIRAPVSGYAGLKVF 298
302 TEGAVVRAAEPLVDIAPDEAFVIOAHFQPTVDVNHGMQVTEVLRPAFHSAGNDPREH 361
299 TDGGVIGPEGLMTVYVPSNDSLEVEGQLAVNLVDRIHSLPEMLFTVANSQKTRVYGE 358
362 DEVAVADRISDPQKQARLFLGIYRVVQKLPRLRG-RVTAQMPAQVIVPTGERTVLOYL 420
359 VTMVADRLDDQGNKOPYALRAQVDAAM-CKLKGLQIRPMANQVFPVTERGSLNLT 417
421 FSPRLDFTLTWRE 434
418 FKPLFDRAHVALAE 431

RESULT 15

S12526
metallopeptidase export system membrane fusion protein prte [validated] - Erwinia chrysanthemi
C:Species: Erwinia chrysanthemi
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 26-May-2000
C:Accession: S12526; T48667
R:Letoffe, S.; Delepelaitre, P.; Wandersman, C.
EMBO J. 9, 1375-1382, 1990
A:Title: Protease secretion by Erwinia chrysanthemi: the specific secretion functions at
A:Reference number: S12524; MUID:90228333; PMID:2184029
A:Accession: S12526
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-448 <LEFT>
A:Cross-references: GB:X53253; NID:941546; PIDN:CMA37343.1; PID:941549
R:Delepelaitre, P.; Wandersman, C.
submitted to the EMBL Data Library, April 1991
A:Description: C-terminal secretion signal in protease secreted by gram negative bacteria
A:Reference number: Z24521
A:Accession: T48667
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-448
A:Cross-references: EMBL:M60395; PIDN:AAA63635.1
A:Experimental source: strain B374
C:Genetics:
A:Gene: prte
C:Complex: the export system consists of the ATP-binding protein prtd (PIR:S12525); the
76141]
C:Function:
A:Description: involved in the export of metalloproteinases [validated, MUID:90228333]
C:Superfamily: hemolysin secretion protein D; lipoyl/biotin-binding homology
C:Keywords: membrane-associated protein; protein export

F:64-107,300-Domain: lipoyl/biotin-binding homology #status atypical.<LPB>

Query Match 23.4%; Score 506.5; DB 2; Length 448;
Best Local Similarity 28.8%; Pred. No. 3.8e-24;
Matches 128; Conservative 93; Mismatches 191; Indels 33; Gaps 7;

8 RPTDNPQAVARIYGIALTFVGLGMAFAPLDSAVIANG-VSAEVSQDVQHEGGL 66
20 RASRDEERLARLRGWMVLVLAFCGFLMLALPLDGVANQGVVVSQNKIKVQHGGIV 79
67 AKIIVREGKKAQGVLELPTQANAAAGITRNOYVALKAMEARLARERDQPSISFPA 126
80 DRIQVKGDRVAAQGVLLTLNVDARTTSEGSGQYDOLIAREARLARERDQPSISFPA 139
127 DLTSGRADPMVARIADDEQAQFTERROTIOGVDMNAQRLQOSEIEGIDRQOGIKQ 186
140 RLTOARQREPMATIALDE-----DLRSQSGKLEIDVRASIDGLETS 185
187 LGFIE-----DELIDRLKLYDKGLVPRRLALEARAGSLSGSIGRLTADR 232
186 LGALQKVMSSKQSEQATLSQGLRLPLAADVYPRNKLTERLFAQVSGSLAQTSGEV 245
233 SKAVQASDPTOLKVRQIKQEFFEQVQSITETRVRLAEVTEKEVVAQAQRIRKIVSPV 292
246 GTRTRDIOQOKLRIARQROEDYKENVSELSDVQAKLNEVYSOREKADFNLANVQRAVPA 305
293 GTRQNLRFPTBGAVVRAAEPLVDIAPDEAFVIOAHFQPTVDVNHGMQVTEVLRPAFHS 352
306 GTVAVMKITFTGEGVIAIPQVMMDIVPEQDPLVDGRIPVEMDKVWSGLPVLQFTAFSQ 365
353 AGNPDEPERRHPVAVVADRISDPQKQARLFLGIYRVVQKLPRLRG-RVTAQMPAQVIVP 410
366 STTPRPVPGVITLLSADRLVD-EKDDPTTYGL-RIVQSEGRKSLGLETIKPMPVQGFPR 423
411 TGERTVLOYLFSPRLDFTLTWRE 435
424 TGERSEFNYLFPKPLMDRMHLALTEE 448

Search completed: April 22, 2003, 12:32:47
Job time : 23 secs

THIS PAGE BLANK (USPTO)

PT polypeptide, useful for treating sewage, waste water and in the pulping process -

XX Claim 1; Page 35; 46pp; English.

CC The invention relates to a Caulobacter host cell, for expression and
CC secretion of a heterologous polypeptide. The Caulobacter host cell
CC comprises at least one surface layer transport protein having an
CC amino acid sequence homologous to the Caulobacter crescentus Rsad or
CC Raaf proteins (AAB22786, AAB22787). The host cell further comprises a
CC DNA construct comprising DNA encoding a polypeptide heterologous to a
CC surface layer (S-layer) protein of the cell 5' from, and operably linked
CC to, a DNA encoding a Caulobacter surface layer protein secretion signal,
CC with the proviso that when the cell comprises transport proteins having
CC the same sequence as both the Rsad and Raaf proteins, the secretion
CC signal is not from the Caulobacter crescentus surface layer protein
CC (Rsaa). The invention also encompasses a method for identifying a
CC Caulobacter suitable for use as a host cell for expression and secretion
CC of a heterologous polypeptide comprising the detection of homologues of
CC the rsad or raaf genes using a rsad/rsaa hybridisation probe. The host
CC cell is used for the expression and secretion of a heterologous
CC polypeptide. The modified Caulobacter cells may be used to treat sewage
CC and waste water. The cells may also be grown in wood pulp suspensions and
CC then used in the wood pulping process. They can also be used in fish
CC vaccines. Caulobacter species are able to form biofilms, and can attach
CC themselves to surfaces without producing the extracellular enzymes or
CC polysaccharide slimes that are characteristic of most other
CC surface-attached bacteria. Caulobacter are therefore particularly suited
CC for use in bioreactor systems. The present sequence represents the
CC Caulobacter crescentus membrane fusion protein (MFP), Raaf. This protein
CC is anchored in the inner membrane of Caulobacter crescentus and is
CC thought to span the periplasm. Together with an outer membrane protein,
CC it forms a channel that extends from the cytoplasm through the two
CC membranes to the outside of the cell.

XX Sequence 435 AA;

SO Query Match 100.0%; Score 2165; DB 21; Length 435;
Best Local Similarity 100.0%; Pred. No. 3, 2e-178;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKPKRIQPTNFOAVARIGYIATLTFVGLGMAAFPLDSAVIANGVSAEVSQYH 60
DB 1 MKPKRIQPTNFOAVARIGYIATLTFVGLGMAAFPLDSAVIANGVSAEVSQYH 60
QY 61 LEGGMLAKIIVREGEKVKAGVLELDPYQANAAAGITRNQYVALKAMEARLLAERDORP 120
DB 61 LEGGMLAKIIVREGEKVKAGVLELDPYQANAAAGITRNQYVALKAMEARLLAERDORP 120
QY 121 SISPADLTSGRADPVARAIADBOAQFTERRQTIQGVDMNAORLOYOSEIEGIDROT 180
DB 121 SISPADLTSGRADPVARAIADBOAQFTERRQTIQGVDMNAORLOYOSEIEGIDROT 180
QY 181 OGLDQGLFEIDELIDRLKLYDKGVPRLALAEARSGISGRTARASRAVQAS 240
DB 181 OGLDQGLFEIDELIDRLKLYDKGVPRLALAEARSGISGRTARASRAVQAS 240
QY 241 DTOLKVRQIKOEFEEYQSOSITETRVRLAEVTEKEVNASDAQRIKIYSPNGTAQNLF 300
DB 241 DTOLKVRQIKOEFEEYQSOSITETRVRLAEVTEKEVNASDAQRIKIYSPNGTAQNLF 300
QY 301 FTBGAIVVAAEPLVDIAPEDEAFVIOAHFOPTDVNVMHGMVTEVRLPAFHSAGNPDEP 360
DB 301 FTBGAIVVAAEPLVDIAPEDEAFVIOAHFOPTDVNVMHGMVTEVRLPAFHSAGNPDEP 360
QY 361 HDPAAVADRISDPQKARLFLGIVRVYKQLPRLRGVATGMPAOVIVPGERIVLOYL 420
DB 361 HDPAAVADRISDPQKARLFLGIVRVYKQLPRLRGVATGMPAOVIVPGERIVLOYL 420
QY 421 FSPPLDRLTTMREE 435
DB 421 FSPPLDRLTTMREE 435

RESULT 2

AAW05298

ID AAW05298 standard; Protein; 443 AA.

XX AC AAW05298;

XX DT 30-DEC-1996 (first entry)

XX DE Esterase secretory protein (ORF2 product).

XX KW Esterase secretory gene; protein secretion.

XX OS Serratia marcescens strain Sr41 (FERM BP-487).

XX PN EP733707-A2.

XX PD 25-SEP-1996.

XX PF 20-MAR-1996; 96EP-0104429.

XX PR 23-MAR-1995; 95JP-0063772.

XX PA (TANA) TANABE SEIYAKU CO.

XX PI Akatsuka H, Kawai E, Shibatani T;

XX DR WPI: 1996-444599/45.

XX DR N-PSDB: AAT39662.

XX PT New isolated esterase secretory gene from Serratia - used to

XX PS Increase the productivity and extracellular secretion of esterases

XX SO Claim 3; Page 21-22; 29pp; English.

CC 3 Different proteins (AAW05297-99) participate in the mechanism of
CC secretion of esterase by Serratia marcescens Sr41. They are
CC encoded by an esterase secretory gene (AAT39662) that includes 3
CC open reading frames. The proteins can be produced by transformed
CC host cells, pref. S. marcescens or E. coli, carrying vector
CC plasmids incorporating the esterase secretory gene. The esterase
CC is secreted from the host cell, and is recovered from the culture
CC medium and from within the cells. The esterase is useful in
CC hydrolysis reactions.

XX Sequence 443 AA;

SO Query Match 24.1%; Score 521.5; DB 17; Length 443;
Best Local Similarity 29.3%; Pred. No. 1, 6e-36;
Matches 133; Conservative 88; Mismatches 194; Indels 39; Gaps 7;

QY 6 IORPTDNF-----QAVARIGYIATLTFVGLGMAAFPLDSAVIANGVSAEVS-Q 56
DB 5 IORPTDNF-----QAVARIGYIATLTFVGLGMAAFPLDSAVIANGVSAEVS-Q 56
QY 57 DVQHEGMLAKIIVREGEKVKAGVLELDPYQANAAAGITRNQYVALKAMEARLLAER 116
DB 57 DVQHEGMLAKIIVREGEKVKAGVLELDPYQANAAAGITRNQYVALKAMEARLLAER 116
QY 117 DQRSISPADLTSGRADPVARAIADBOAQFTERRQTIQGVDMNAORLOYOSEIEG 176
DB 117 DQRSISPADLTSGRADPVARAIADBOAQFTERRQTIQGVDMNAORLOYOSEIEG 176
QY 176 DQRSISPADLTSGRADPVARAIADBOAQFTERRQTIQGVDMNAORLOYOSEIEG 176
DB 176 DQRSISPADLTSGRADPVARAIADBOAQFTERRQTIQGVDMNAORLOYOSEIEG 176
QY 177 DRQTOGLKID-----OLGFEDELIDRLKLYDKGVPRLALAEARSGISGRTAR 229
DB 177 DRQTOGLKID-----OLGFEDELIDRLKLYDKGVPRLALAEARSGISGRTAR 229
QY 229 DRQTOGLKID-----OLGFEDELIDRLKLYDKGVPRLALAEARSGISGRTAR 229
DB 229 DRQTOGLKID-----OLGFEDELIDRLKLYDKGVPRLALAEARSGISGRTAR 229
QY 230 ADRSKAVQASDTOLKVRQIKOEFEEYQSOSITETRVRLAEVTEKEVNASDAQRIKIY 289
DB 230 ADRSKAVQASDTOLKVRQIKOEFEEYQSOSITETRVRLAEVTEKEVNASDAQRIKIY 289
QY 289 ADRSKAVQASDTOLKVRQIKOEFEEYQSOSITETRVRLAEVTEKEVNASDAQRIKIY 289
DB 289 ADRSKAVQASDTOLKVRQIKOEFEEYQSOSITETRVRLAEVTEKEVNASDAQRIKIY 289
QY 297 GRIGLOKQOLLESQQRIDQREFADYQREVRYQLAQTDASDFRKKLQMADEDLGNTAITS 297
DB 297 GRIGLOKQOLLESQQRIDQREFADYQREVRYQLAQTDASDFRKKLQMADEDLGNTAITS 297
QY 349 PVNGTAONLREFTGAVVAAEPLVDIAPEDEAFVIOAHFOPTDVNVMHGMVTEVRLPA 349
DB 349 PVNGTAONLREFTGAVVAAEPLVDIAPEDEAFVIOAHFOPTDVNVMHGMVTEVRLPA 349

```

Db      PVDSGVVGLNFTGGVGVAGGDHLMADVPSQATLVYDSRLKVDLEKVVYNGLEPVLDMETA 357
Qy      FHSAGNDPRERHDVNAVDRSD-----POKORALEL-----GYRVYKQALPRHLRGVRA 401
Db      FNOKTRFIPCTYVLVNSDRLYVDANCEPRTYOMQVYSPSGMKMLSGEDIKP----- 409
Qy      GMPAOVIVPTGERTVLOYSPLDLRTTWREE 435
Db      GMPVEVFPVTKTSRLSLSTLEKPIIDRAHTSITJEE 443

```

RESULT 3
AAY55920
ID AAY55920 standard; Protein; 444 AA.

AC AAY55920;

DT 15-FEB-2000 (first entry)

Pseudomonas fluorescens ABC transporter cassette component B.

KM ATP-binding cassette; transporter; operon; *LipBCD*; *Serratia marcescens*;
KM microbe; protein secretion.

Pseudomonas fluorescens.

PN JP11276172-A.

PD 12-OCT-1999.

PF 27-MAR-1998; 98JP-0080597.

PR 27-MAR-1998; 98JP-0080597.

PA (TANA) TANABE SEIYAKU CO.

DR WPI: 1999-626936/54.

XX
DE

PT secrete proteins
vv

PS Claim 2; page 16-17; 28pp; Japanese.
xx

CC This sequence represents an Atr-binding cassette (ABC) transporter
CC component B from an *Pseudomonas fluorescens* strain 33 and encoded by
CC an ABC transporter operon. The operon comprises 3 genes where the
CC termination codon of the first gene overlaps with the initiation codon
CC of the second gene. The sequences of the encoded proteins have 60, 44
CC and 46% homology respectively to the LibCD proteins from *Serratia*
CC marcescens. The novel gene and protein can give or increase the
CC ability of a microbe for secreting a protein.

Sequence 444 AA;

Query Match	22.9%;	Score 495;	DB 20;	Length 444;
-------------	--------	------------	--------	-------------

Matches 134; Conservative 86; Mismatches 201; Indels 12; Gaps 7;

```

QY      9 PTDNFQAVARIGCYITALTVEGLG---WAAFAPLDSAVIANG-VWSAEVSQDVHLEGG 64
      | : ||: | | | | | : ||: | : | : | : | | | |
DB     18 PEHGAREFAIRGWM--LTVVGAGGCFLLMASLAPLDQIGIPQVGTVVVSGKRAAYQTLSPG 74

```

```
0y      65 MIAKILVRESEKVKAGOVLEFDPTQANAAAGITRNQYALKAMEARLLAERDQRPSISF 124
        ::::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      75 VNSILVREEAVKKGQPLFRIDQTQNADVHSLQAQIYNAMASVAARKWSEPDNOSTTF 134
```

Qy 125 PADITSGRADPMVARIADEQAOTERRQTIOGVDLMAOQLYQSEITEGIDRQTGLK 184
||| | : : : : ||| : : : :
Db 135 PAEL-SGNPDQALALVLEGORQLFSSREAFARPOAGIRANIEGATQNLNGRRARSDLT 193

185 DQGFIEDIDLKRLYDKGLVPRPRLIALEARAGSLSGSIGRLTADRSKAVQASPTOL 244

Db	194	AQNSLRQJLNNLQPLADNGIIPNNRLMEYOROLSOYQODLAQNTGSSGVEGJILESLR	253
OY	245	KVRQIKQEFFEVSOSITETRVRLAEYTEKEVVASDAOKRIKIVSPVNGTAONLREFTEG	304
		: : : : : : : : : :	
Db	254	KLOOHSSEYQKEVPSQLDADQLRSJLTLEOQLTSAGFDLQHSSEIINAPADGIAVNLVSFTEG	313
		: : : : : : : : : :	
OY	305	AVVRAPELVYIAPDEEFAVIOAFOPFDVONVMGWAETEVRLEPAPASAGNPDEERHDV	364
		: : : : : : : : : :	
Db	314	AVVRAGEJLLEIVQDTLVEGRPLPHLVDKVTHLPLDILFAFNQSRPRVPGVSL	373
		: : : : : : : : : :	
OY	365	AVADRI SDPQOARLELGIYRVYKQLE - PHLGRV - TAGNPAQIVPTGERTVLYLFS	422
		: : : : : : : : : :	
Db	374	ISADQMDEKGMPPY - VLRFTVSSSALEKJHLVIKPGMAPEMFIRTGERSLNLYFK	431
		: : : : : : : : : :	
OY	423	PLDRLTRTMMREE	435
		: : : : : : : : : :	
Db	432	PLUDRAGSALTEE	444

RESULT 4	
AAV55923	
ID	AAV55923 standard; Protein; 443 AA

AC AAY55923

DT 15-FEB-2000 (first entry)

DE *Pseudomonas fluorescens* ABC transporter cassette II component B.

KW ATP-binding cassette; transporter; operon; *LipBCD*; *Serratia marcescens*;
KW microbe; protein secretion.

Pseudomonas fluorescens.

PN JP11276172-A.

PD 12-OCT-1999.

PF 27-MAR-1998; 98JP-0080597.

PR 27-MAR-1998; 98JP-0080597.

PA (TANA) TANABE SEIYAKU CO.

DR WPI; 1999-626936/54.

XX

PT secrete proteins

PS Claim 8; Page 23-24; 28pp; Japanese.

CC This sequence represents an ATP-binding cassette (ABC) transporter
CC component B from an *Pseudomonas fluorescens* strain 33 and encoded by
CC the ABC transporter operon II. The operon comprises 3 genes where the
CC termination codon of the first gene overlaps with the initiation codon
CC of the second gene. The sequences of the encoded proteins have 63, 56
CC and 594 homology respectively to the lipase proteins from *Serratia*
CC marcescens. The novel gene and protein can give or increase the
CC ability of a microbe for secreting a protein.

Sequence 443 AA;

Query Match	22.1%	Score 479.5	DB 20	Length 443
Best Local Similarity	29.0%	Pred. No. 7.6e-33		
Matches 130; Conservative	83;	Mismatches 196;	Indels 39;	Gaps 7;

Qy 9 PTDFEQAV-----ARIGYGIALLTFVGLLGWAAPLDSAVIANG-VVSAEVSQDVQHLE 62
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 14 PVNNVLALDDKKYARLGLWLVLCGFAGFLGMAALPLDKGAVASGKWWVGSHRKTVOHPS 73

07 63 GGM LAKILVREGGEKVAGQVLELDPTQANNAAGITRNOYVALKAMEARLLAERDQPSI 122

PR 09-OCT-1998; 980S-0103796.
 PR 25-FEB-1999; 990S-0121528.
 XX
 PA (CHIR) CHIRON CORP.
 XX (GENO-) INST GENOMIC RES.
 XX
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
 PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;
 XX
 DR WPI: 2000-062150/05.
 XX N-PSDB: AA254335.
 XX
 PT Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics
 XX
 PS Claim 2: Page 1243; 1453pp; English.
 XX
 AA253015 to AA254536, AA254577 to AA254615, and AA254253 to AA25941
 CC represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides
 CC and polypeptides. AA254537 to AA254576 and AA254616 to AA25473 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the
 CC presence of Neisseria bacteria, or to raise antibodies. They may also
 CC be used to screen for agonists or antagonists, which may themselves
 CC have use as antibacterial agents. The polynucleotides of the invention
 CC may also be used in gene therapy protocols.
 CC
 XX
 XX
 SQ Sequence 477 AA;

Query Match 15.2%; Score 329; DB 21; Length 475;
 Best Local Similarity 26.2%; Pred. No. 8e-20;
 Matches 128; Conservative 78; Mismatches 188; Indels 94; Gaps 18;

OY 1 MKPPIQRTDNE-----QAVARIGYGIATFVGLGMAFAPLDSAV 44
 DB 27 LKPRRTAEBOAFPAHLELDTFVSAAPKMAAF--IMAFALIAL-WSWFGKIDIVA 82
 OY 45 IANG-VVSAEVSODYOHLGEGMLAKTLVREGKVRAGOVLEFELPT-----QANA 93
 DB 83 AASGTVSGGSRKTIQPLETAVKAVHRDQGVKQGETIALELVGDSDDVQSEQALQ 142
 OY 94 AAGITRNQYVA-LKAMEARLAEKDQPSISFPADLTQORADPMVRAIADQAOFTERR 152
 DB 143 AAOISKLRYEAVLALLESR-----TVPHIDMAQARSIGLSADADVQSQ 185
 OY 153 QTIQGVOLMNAQRLOYSEIEGIDRQGLKQ-----LGFIDE-LIDIRKLYDGL 205
 DB 186 VLAQHQYQAMAAQDQALQDQALSGHQAELQSAQAQOKIVSGAIEHQQTADRRRLADNF 245
 OY 206 VPRRLALEARAGS-----LSGSIGRLTADRSKAVQASDQOLKROIKOEFPE---QVS 258
 DB 246 ISEHAFLEQGSVSNMMDLESTRQGMQIQALIAQACNRNLNQNKLROTLDLRLQAN 305
 OY 259 QSITEYRLAEVTEKEVYASDAQKRIIVSFPVNGTAONLRFETGCAVYRAAEPLVDIAP 318
 DB 306 EOIDQYRGO---TDK---AKORQOOLMTIQSPADGTVELATYIVGVQVQAOKMAMVAP 358
 OY 319 EDEAVFVIAHQPTQDVVDVHMGVTEVRLPAF-----HSGNPNPERRHDPVAVADRISL 372
 DB 359 DDDKKADVVLANKDIGVEQOGDAVVKIESPPYRIGLIGKVKVSVSHDAS----- 411
 OY 373 POKQARLEAGIVRVNDVKQLPPH--LRGR--VTAGMPAOVIVPGEFVTLQYLFSPULD 426
 DB 412 -HEQ---LGLVYTAIVSLDKHTLINDGKAVVLFGAMNVVLEIKRKRRVLDVLLSPLOT 466
 OY 427 TLRTTMR 434
 DB 467 RLDESFR 474

RESULT 7
 ID AAM22161
 XX AAM22161 standard; Protein; 477 AA.
 XX
 AC AAM22161;
 XX
 DT 16-FEB-1998 (first entry)
 XX
 DE ApxIIID protein.
 XX
 KW RTX toxin; apxICA gene; apxIIB gene; apxIIABCD gene;
 KW repeat in toxins toxin; cell-associated RTX toxin; vaccine production;
 KW therapy; A. pleuropneumoniae infection; swine pleuropneumonia.
 XX
 OS Actinobacillus pleuropneumoniae.
 XX
 PN CA2170839-A.
 XX
 PD 02-SEP-1996.
 XX
 PF 01-MAR-1996; 96CA-2170839.
 XX
 PR 01-MAR-1995; 95US-0396244.
 XX
 PA (UYGU-) UNIV GUELPH.
 XX
 PI MacInnes J, Mallard B, Ricciattl P, Rosendal S;
 XX
 DR WPI: 1997-245536/23.
 DR N-PSDB: AAT73220.
 XX
 PT Preparations of microorganisms producing cell-associated RTX toxins
 PT - especially for production of vaccines against swine
 PT pleuro-pneumonia
 XX
 PS Disclosure; Pages 113-114; 151pp; English.
 XX
 CC AAM22151-W22161 represent A. pleuropneumoniae RTX (repeat in toxins)
 CC toxins. These sequence are encoded by the apxICA, apxIIB, apxIIABCD, and
 CC apxIIABCD genes (see AAT73217-T73220), and can be expressed by
 CC microorganisms used in the preparations of the invention. The
 CC preparations are bacterial preparations comprising one or more isolated
 CC and purified strains of a microorganism that produces one or more RTX
 CC toxins, where the strains have at least one cell-associated RTX toxin.
 CC The preparations are used for production of vaccines for the prophylaxis
 CC and treatment of infectious diseases caused by microorganisms that
 CC produce RTX toxins, where the strains have been attenuated or
 CC inactivated. The vaccines are preferably against Actinobacillus
 CC pleuropneumoniae infection (swine pleuropneumonia). It has been found
 CC that A. pleuropneumonia produces significant quantities of
 CC cell-associated RTX toxins when cultured under certain conditions, and
 CC that the whole-cell protein composition of the cultures corresponds to
 CC the whole-cell protein profiles obtained from cells recovered at
 CC necropsy from the pleural fluid of infected swine. Vaccination with a
 CC bacterin prepared from heat-inactivated cultures having significant
 CC quantities of cell-associated RTX toxins give significant protection of
 CC swine against challenge with homologous strains.
 XX
 SQ Sequence 477 AA;
 OY
 Query Match 14.5%; Score 313.5; DB 18; Length 477;
 Best Local Similarity 25.3%; Pred. No. 1.8e-18;
 Matches 112; Conservative 88; Mismatches 190; Indels 53; Gaps 13;
 OY 19 IGYGIATLFYVGLGMAFAPLDSAVIANG-VVSAEVSODYOHLGEGMLAKTLVREGKRV 77
 DB 60 IAYLIMLFLFLAIV-ISIISKVEIYASATGKLVFSGHSKEIKPIENALVKDIFVADGQFV 118
 OY 78 KAGQVLEL-----DPLQANAAAGITRNQYVALKAMEARLAEKDQPSISF-PADIT 129
 DB 119 EKGQILNLTALGCDADQKOTKVSGLERLDGYRKYSLYSI--EHNRLPLDFFNOADFD 176

QY 130 S-QRADPVARAIAD-----QAQFERRQTIGQVYDLMNAQRLQYQSEIEGIDR 178
 DB 177 SVOEDKTKGARHLITEQETWOKRYOKELAVQROKAEQVLAIRKYESASRIE---- 232
 QY 179 QTOGAKDQGLFTEDELIDRLKLYDKGLVPRRLALLEAR-----ASLSGSGIGRLTADRSK 234
 DB 233 -----KELSLDKRLKYDVKSISKHELLAQENRYEASNELSVYQSHKEVESD 280
 QY 235 AVQASDPTQLKVRQIKQEFQVQSITETRYRLAEVTEKEVYASDAQRIK--IVSPVN 292
 DB 281 LKQAEDELKLVTLQFKSDILEKIQNIREKQLTLELEKNE-----QROLASITIRAPVS 334
 QY 233 GTAQNLREFTBGAIVRAAEPLVDIAPDEBAFYQAHFOPTVDVNHGMVTEVRLPAFHS 352
 DB 335 GTVQDLKTRHKGAVTTAETLAVIAPEDDVEVSALIONKDVGEVIEQEAIVKEVTEFPY 394
 QY 353 AGNDPERHDPVAVADRISDPQKARFLGIVRVKOLPRHLRG-RVTAGMPAQVIPT 411
 DB 395 TRYGLYGRKVTITLDAIEHPQ-LGLVFNSTIEINKKTLTGDKETIQLGSGMSVIAETKT 453
 QY 412 GERVLYQTLFSPRLDRLTMTRE 434
 DB 454 GERSVIFSLSPLESITESLRE 476

RESULT 8
 AAY51414
 ID AAY51414 standard; protein; 477 AA.

AC AAY51414;

DT 05-MAY-2000 (first entry)

DE A. pleuropneumoniae aprXIID protein.

XX RTX toxin; vaccine; Repeats in Toxins; immunostimulatory; antimicrobial;
 KW antiinflammatory; antirheumatic; antidiarrhoeal; treatment; pneumonia;
 KW pleuropneumonia; septicemia; nephritis; arthritis; endocarditis;
 KW shipping fever; abortion; whooping cough; sleepy foal disease; joint ill;
 KW urinary infection; peritonitis; meningitis; gastroenteritis;
 KW passive immunization; aprXIID.

XX Actinobacillus pleuropneumoniae.

XX US6019984-A.

XX 01-FEB-2000.

XX 23-DEC-1996; 96US-0772270.

XX 01-MAR-1995; 95US-0396244.

XX (UVGU-) UNIV GUELPH.

XX Mallard B, Rosendal S, MacInnes J, Ricciattl P;

XX WPI; 2000-146864/13.

XX N-PSDB; AAZ88587.

PT Bacterial preparation comprising microorganisms which produce a member
 PT of the Repeats in Toxins (RTX) family, useful for treating swine
 PT pleuropneumonia, arthritis in swine, shipping fever and abortion in
 PT cattle, and sleepy foal disease -

XX Disclosure; Column 89-98; 96pp; English.

XX This invention describes a novel bacterial preparation (I) which
 CC comprises one or more isolated and purified strain(s) of a microorganism,
 CC cultured in tryptone yeast extract (TYE) broth, which produces one or
 CC more RTX toxins (belonging to the family of toxins referred to as Repeats
 CC in Toxins), where the strain(s) have at least one RTX toxin which is
 CC cell-associated. The products of the invention have immunostimulatory,

CC antimicrobial, antiinflammatory, antirheumatic and antidiarrhoeal activity.
 CC The bacterial preparation may be used as vaccines for the prophylaxis and
 CC treatment of infectious diseases caused by strains of microorganisms
 CC which produce one or more RTX toxins. The infectious diseases are swine
 CC pleuropneumonia, pneumonia, septicemia, nephritis and arthritis in swine;
 CC septicemia, nephritis, endocarditis and arthritis in piglets; shipping
 CC fever and abortion in cattle; whooping cough, sleepy foal disease or
 CC joint ill (purulent nephritis, arthritis) in foals; septicemia,
 CC polyarthritis and abortion in horses; and urinary infections,
 CC peritonitis, meningitis, and gastroenteritis. The bacterial preparations
 CC may also be used to prepare antibodies which may be used as a means of
 CC passive immunization. This sequence represents the Actinobacillus
 CC pleuropneumoniae aprXIID protein described in the method of the
 CC invention.

CC Sequence 477 AA;

Query Match 14.5%; Score 313.5; DB 21; Length 477;
 Best Local Similarity 25.3%; Pred. No. 1.8e-18;
 Matches 112; Conservative 88; Mismatches 190; Indels 53; Gaps 13;

QY 19 IGVGIIATFVGLGMAFAPLDSAVIANG-VSAEVSQDVYHLEGMLAKIIVEGKV 77

DB 60 IAYLMLFLFLAIY-ISIISVEIYASATKIVFSGHKEIKPIENALVKDIPVQGGPV 118

QY 78 KAGQVLFEL-----DPTQANAAGITRNOYVALKAMEARLAEKDPSPISF-PADLT 129

DB 119 EKQGLLLNLTLGCDADKQKTVSLGLERLDGYRYKSLYSI--EHNRLPLDFNQADFD 176

QY 130 S-QRADPVARAIAD-----QAQFERRQTIGQVYDLMNAQRLQYQSEIEGIDR 178

DB 177 SVOEDKTKGARHLITEQETWOKRYOKELAVQROKAEQVLAIRKYESASRIE---- 232

QY 179 QTOGAKDQGLFTEDELIDRLKLYDKGLVPRRLALLEAR-----ASLSGSGIGRLTADRSK 234

DB 233 -----KELSLDKRLKYDVKSISKHELLAQENRYEASNELSVYQSHKEVESD 280

QY 235 AVQASDPTQLKVRQIKQEFQVQSITETRYRLAEVTEKEVYASDAQRIK--IVSPVN 292

DB 281 LKQAEDELKLVTLQFKSDILEKIQNIREKQLTLELEKNE-----QROLASITIRAPVS 334

QY 293 GTAQNLREFTBGAIVRAAEPLVDIAPDEBAFYQAHFOPTVDVNHGMVTEVRLPAFHS 352

DB 335 GTVQDLKTRHKGAVTTAETLAVIAPEDDVEVSALIONKDVGEVIEQEAIVKEVTEFPY 394

QY 353 AGNDPERHDPVAVADRISDPQKARFLGIVRVKOLPRHLRG-RVTAGMPAQVIPT 411

DB 395 TRYGLYGRKVTITLDAIEHPQ-LGLVFNSTIEINKKTLTGDKETIQLGSGMSVIAETKT 453

QY 412 GERVLYQTLFSPRLDRLTMTRE 434

DB 454 GERSVIFSLSPLESITESLRE 476

RESULT 9

AAW22154
 ID AAW22154 standard; Protein; 478 AA.

AC AAW22154;

DT 16-FEB-1998 (first entry)

DE AprXID protein.

XX RTX toxin; aprXICA gene; aprXIBD gene; aprXIAB-C gene; aprXIITABCD gene;
 KW repeat in toxins toxin; cell-associated RTX toxin; vaccine production;
 KW therapy; A. pleuropneumoniae infection; swine pleuropneumonia.

XX Actinobacillus pleuropneumoniae.

XX CA2170839-A.

XX 02-SEP-1996.

[illegible]

0y 399 --VTAGPAAOIVPTGERTVLOVFSPRLDPLRTTME 434
 ::| | :| | | | :| | | | :: :| |
 Db 439 EIEIGSGMSVTAETIKTGERSVIYSTLSPLEESVSLEK 477
 RESULT 10
 ID AAY51408 standard; protein: 478 AA.
 AAY51408
 AAY51408;
 DE 05-MAY-2000 (first entry)
 XX
 DE A. pleuropneumoniae clydID protein.
 XX
 XX RTX toxin; vaccine; Repeats in Toxins; immunostimulatory; antimicrobial;
 KW antiinflammatory; antiarthritic; antilaborative; treatment; pneumonia;
 KW pleuropneumonia; septicemia; nephritis; arthritis; endocarditis;
 KW shipping fever; abortion; whooping cough; sleepy foal disease; joint ill;
 KW urinary infection; peritonitis; meningitis; gastroenteritis;
 XX passive immunization; clydID.
 XX Actinobacillus pleuropneumoniae.
 OS US6019984-A.
 XX
 PN 01-FEB-2000.
 PD
 PF 23-DEC-1996; 96US-0772270.
 XX
 PR 01-MAR-1995; 95US-0396244.
 XX
 PA (UYGU-) UNIV GUELPH.
 XX
 PI Mallard B, Rosendal S, MacInnes J, Ricciattl P:
 DR WPI: 2000-146864/13.
 DR N-PSDB: AAZ88585.
 XX
 PT Bacterial preparation comprising microorganisms which produce a member
 PT of the Repeats in Toxins (RTX) family, useful for treating swine
 PT pleuropneumonia, arthritis in swine, shipping fever and abortion in
 PT cattle, and sleepy foal disease -
 XX
 PS Disclosure: Column 59-62; 96pp: English.
 XX
 CC This invention describes a novel bacterial preparation (I) which
 CC comprises one or more isolated and purified strain(s) of a microorganism,
 CC cultured in tryptone yeast extract (TYE) broth, which produces one or
 CC more RTX toxins (belonging to the family of toxins referred to as Repeats
 CC in Toxins), where the strain(s) have at least one RTX toxin which is
 CC cell-associated. The products of the invention have immunostimulatory,
 CC antimicrobial, antiinflammatory, antiarthritic and antilaborative activity.
 CC The bacterial preparation may be used as vaccines for the prophylaxis and
 CC treatment of infectious diseases caused by strains of microorganisms
 CC which produce one or more RTX toxins. The infectious diseases are swine
 CC pleuropneumonia, pneumonia, septicemia, nephritis and arthritis in swine;
 CC septicemia, nephritis, endocarditis and arthritis in piglets; shipping
 CC fever and abortion in cattle; whooping cough, sleepy foal disease or
 CC joint ill (purulent nephritis, arthritis) in foals; septicemia,
 CC polyarthritis and abortion in horses; and urinary infections,
 CC peritonitis, meningitis, and gastroenteritis. The bacterial preparations
 CC may also be used to prepare antibodies which may be used as a means of
 CC passive immunization. This sequence represents the Actinobacillus
 CC pleuropneumoniae clydID protein described in the method of the invention.
 CC
 SO Sequence 478 AA:
 Query Match 12.1%; Score 262; DB 21; Length 478;
 Best Local Similarity 24.2%; Pred. No. 4.9e-14;
 Matches 111; Conservative 86; Mismatches 178; Indels 84; Gaps 18;

```

Db 60 IAYLMFLFLALV-ISTVSHVEIYATATGKIASDSRKEKPIENALVKELFVODGQFV 118
QY 78 KAGOVLELDEPTQANNAAGITRNOYVALKAMEAR--LLAE--RDQRSISEPAD--LTS 130
Db 119 EKDOLLHLTALGADADQDKTSSLSLTKLERYREILLEVAADRLLIETTKDFKHA 178
QY 131 QRADPMVARAIADDEQAOTER-----KQTQGVVDLMAQRLOYSIESIDRQT 180
Db 179 TEEDTRIRYLTTEQEFEMOKOKYOKELALORREAEKQTVLANIRK-----YEGISR-- 230
QY 181 QGLKDLQLEFIDE-LIDRKLYDKGLVPRPLLALEAR---AGSLSGSIGRLTRADRSKA 235
Db 231 -----VENERLKDKLKFNSKSTSKHDVLTOENHIEAVNLAYKSKSLNVESDL 281
QY 236 VQASDQLKVRQIKQEFEEQVSSITETRVRLAEVTEKEVYASDAQRIKIV--SPYNG 293
Db 282 RQAKEEHLITQLFRADILEKIKQNV-EAKQSLSELEK-----NEQKQIASVIRAPVSG 335
QY 294 TAONLRFETEGAVYRAAEPLVDIAPEDAEFYIOAHFQPTDQVNVHMGVTEVRLPAFISA 353
Db 336 TVQQLKHTHTVGAVVTTAETLWVIAEDVDLEVTALIQNKDGIEFGODAVIKVETTF-- 392
QY 354 GNPPPERHDPVA-----VADRIKDPOKQARFLGIVRVYDVKQLPRHLRGR--- 398
Db 393 -----PYTRGYLGMKVKKNITLLEAIEHPQ--LGLVFNSLISIDRKT-----LSGKDGK 438
QY 399 ---VTAGNPAOVIVPTGERIVLQYLFSPLRDLTFTTME 434
Db 439 EIEIGSGMSVTAETKGTGRSVISYLLSPLESVSSESLRE 477

```

RESULT 11

AAV75574
ID AAV75574 standard; protein; 435 AA.

AAV75574;

21-MAR-2000 (first entry)

Neisseria meningitidis ORF 764 protein sequence SEQ ID NO:2622.

Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
antibacterial; gene therapy.

Neisseria meningitidis.

MO957280-A2.

11-NOV-1999.

30-APR-1999; 99WO-US09346.

01-MAY-1998; 98US-0083758.

31-JUL-1998; 98US-0094869.

02-SEP-1998; 98US-0098994.

02-SEP-1998; 98US-0099062.

09-OCT-1998; 98US-0103749.

09-OCT-1998; 98US-0103794.

09-OCT-1998; 98US-0103796.

25-FEB-1999; 99US-0121528.

(CHIR) CHIRON CORP.

(GENO-) INST GENOMIC RES.

Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M,
Petersen J, Piza M, Rappoli R, Ratti G, Scalato E, Scarselli M,
Tetelin H, Venter JC,
WPI; 2000-062150/05.
N-PSDB; AA254336.

PT Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics
Claim 2; Page 1244; 1453pp; English.
AA53015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941
represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
and polypeptides. AA254537 to AA254576 and AA254616 to AA25473 represent
PCR primers used in the exemplification of the present invention. The
polypeptides, the polynucleotides, antibodies and compositions of
the invention can be used as vaccines, as diagnostic reagents, and as
immunogenic compositions. The polypeptides can be used in the
manufacture of medicaments for treating or preventing infection due to
Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
presence of Neisseria bacteria, or to raise antibodies. They may also
be used to screen for agonists or antagonists, which may themselves
have use as antibacterial agents. The polynucleotides of the invention
may also be used in gene therapy protocols.

Sequence 435 AA;

Query Match 12.1%; Score 261.5; DB 21; Length 435;

Best Local Similarity 25.2%; Pred. No. 4.7e-14;
Matches 102; Conservative 70; Mismatches 163; Indels 69; Gaps 14;

```

QY 23 IIALTFGLGMAFAPLDSAVIANG-VSAEVSODVOHLEGMLAKLVREGKVRAGQ 81
Db 62 IMAFALLALIL-MSWFGKIDIVAAASGKTVSGRSKTIQPLETVVAVKHVHDGQHVXGE 120
QY 82 VLFEIDPT-----QANNAAGITRNOYVA-LKAMEARLAEHRQRPISPADLTS 130
Db 121 TLAELEAVGTSDDVYOSQALQAOALSRLRYEVALALESR-----TV 163
QY 131 QRADPMVARAIADDEQAOTERRQTOGVVDLMAQRLOYSIESIDRQTLGKQD--- 186
Db 164 PHDMAQARSLGSDADQVQSAQVLAHQYQAMAQDQALGQALGHQAELOASAKAQEQKL 223
QY 187 --LGFIDE-LIDRKLYDKGLVPRPLLALEARAGS-----LSSISGRLTRADRSKA 239
Db 224 VSGAIEQOKTADYRRLRADNFIEBHAELEQOOSKSVSNWMDLESTRGMRQIOAALQAE 283
QY 240 SDTOLKVRQIKQEFEE--QVSQSITEFTRVRLAEVTEKEVYASDAQRIKIVSVNGTAQ 296
Db 284 QNRVLNQNLRDLIDLALQANEOIDQYRQ---TDR---AKORQQLMTQSPADGTVO 336
QY 297 NLREFTGAVYRAAEPLVDIAPEDAEFYIOAHFQPTDQVNVHMGVTEVRLPAF----- 350
Db 337 ELATYTVGAVVQAQKMMVVAPODDKMDVEVLVLNKDIGFVEGQGDVAVKIESPPYTRYG 396
QY 351 HSAGNPPDERHDPVAVADRIDSPQKQARFLGIVRVYDVKQLPRH 394
Db 397 YLTGKVKVSVSHDAVS-----HEQ-----LGLVYTVAVVSLDKH 428

```

RESULT 12

AAV24069
ID AAV24069 standard; peptide; 425 AA.

AAV24069;

09-SEP-1999 (first entry)

Salmonella typhimurium peptide sequence SEQ ID NO:38.

Salmonella; Salmonellosis; detection; diagnosis; infection; food.

Salmonella typhimurium.

US5925522-A.

20-JUL-1999.

09-MAY-1997; 97US-0853659.

XX 09-MAY-1997; 97US-0853659.
 XX (BAT) BATTELLE MEMORIAL INST.
 XX Saffer JD, Wong K.
 XX MPI: 1999-418268/35.
 XX N-PSDB: AAX83950.
 XX
 XX Detecting *Salmonella* in humans, animals and food
 XX
 XX Disclosure; Column 161-164; 144pp; English.
 XX
 XX The present sequence represents a *Salmonella typhimurium* peptide
 XX sequence. The *Salmonella* nucleotide (AAX83935), its complement
 XX (AAX83936) and fragments (AAX83937 to AAX83971), are useful for
 XX detecting and diagnosing *Salmonella* infection in humans, animals and
 XX food. The nucleotides will detect many, if not all *Salmonella* species
 XX especially *Salmonella dublin*, *S. enteritidis*, *S. gallinarum*, *S.*
 XX *minnesota*, *S. paratyphi* (types A, B and C), *S. pullorum*, *S. typh ty21a*
 XX and *S. typhimurium*. The fragments may be used to construct a DNA chip,
 XX useful for the simple and inexpensive testing of substances for the
 XX presence of *Salmonella*.
 XX

Sequence 425 AA;

Query Match 9.3%; Score 201.5; DB 20; Length 425;

Best Local Similarity 21.0%; Pred. No. 6; 8e-09;

Matches 94; Conservative 93; Mismatches 187; Indels 73; Gaps 19;

QY 7 GRPTNFQAVARIIGYIALTFVGLGMAFAPLDSAVANGVSA-EVSODVOHLEGGM 65
 DB 3 RQSDHMMI-----IISLTL-IILTFEINSVHGQGVITTKMAQLISLSKGT 55
 QY 66 LAKILVREGEKVKAGVLT--FELDPTQANAAAGITRNOYVALKAMEARLAEERDQPSI 122
 DB 56 IODIYVABEDVFKGELLAKVNLD-----LQKEYGRYRQKGYL--DKDVN-EI 102
 QY 123 SPFADLTSGRADPMV--AAINDEQAQFTREROTIQQVDLNNARLQYQSEIEGIDROT 180
 DB 103 SFLLKENESGLITLDGTRLSLSNKEVK--ANIELVHSQIRAKELKTSLSDEISGLQEK 160
 QY 181 QGLKDLQGFIEDELIDRLKLYDKGLVP-----RPLRLALBARAGSLSGISGRLTADRS 233
 DB 161 SSKREKELLALAEINLSLVKKGISPTYNFLNKKQAYTKVASEINDISSI-TLKDDI 219
 QY 234 KAVQASDTQLKVRQIKQEFPEVOSOSITETRV--RLAEVTEKEVVASDAQRIKIVSPV 291
 DB 220 ELV--VNDIEALNNELRLSLKISKNLOELEVNSTLKVIEKQINEED-----IYSPV 271
 QY 292 NCTAO--NLRFTEGAVVRAEPLVDIAPEDEAFVIOAHFOPTDVNVHMGVTEVRLA 349
 DB 272 DQYIKKINSATTHGVIOADLLEIKPKYRMLADAVKILPKYRQIYDEAVKLDVOS 331
 QY 350 FHSAGNPDERHDPVAVADRISDPQK-----ARLFLGIYRVDKQ-----LPHLHG 397
 DB 332 I-----IQKIKSYNANTINISPDSEYENTGGTIQRYKVIIAIFDVNEDDLMLRK----- 382
 QY 398 RVTAGMPAOVIYPTGERIVLYLFSPL 424
 DB 383 ----GMYTDAVSYTGRKHSIMEYLSPL 405

RESULT 13

AAG67295 ID AAG67295 standard; Protein; 543 AA.

AC AAG67295;

DT 13-NOV-2001 (first entry)

XX Amino acid sequence of ABC membrane fusion protein of avirax21.

XX Avirulence gene; Xa21 resistance gene; avirax21; type I secretion system;
 KW pathogen infection; ABC membrane fusion protein.
 XX Xanthomonas oryzae.
 XX WO200162896-A2.
 XX 30-AUG-2001.
 XX 08-FEB-2001; 2001WO-US04068.
 XX 22-FEB-2000; 2000US-0510206.
 XX (REGC) UNIV CALIFORNIA.
 XX Ronald P, Shen Y, Da Silva FG;
 XX MPI: 2001-514826/56.
 XX N-PSDB: AAH77875.
 XX
 XX Isolated polypeptide of an avirulence gene determinant specific for the
 XX Xa21 resistance gene of plants is useful as a probe or ligand to
 XX identify binding motifs and to develop a system to generate resistance
 XX responses during infection -
 XX
 XX Disclosure; Page 34-36; 47pp; English.
 XX
 XX The present sequence represents an ABC membrane fusion protein. The open
 XX reading frame is derived from a 9 kb avirulence gene determinant
 XX specific for the Xa21 resistance gene from *Xanthomonas oryzae* pv. -
 XX *oryzae*, designated avirax21. Several of the open reading frames of
 XX avirax21 encode predicted proteins with homology to the type I secretion
 XX system of prokaryotes. Avirax21 is secreted by avirulent *Xanthomonas*
 XX *oryzae* pv. *oryzae* (Xoo) bacterial pathogens, and then interacts with
 XX Xa21 gene products in the plant. The avirax21 polynucleotide may be used
 XX as a source of probes or ligands to identify the binding motifs of the
 XX corresponding Xa21 R gene product. The avirax21 polynucleotides and
 XX polypeptides may also be useful in developing a two-component system to
 XX generate resistance responses in plants in response to pathogen
 XX infection.
 XX

Sequence 543 AA;

Query Match 6.9%; Score 149.5; DB 22; Length 543;

Best Local Similarity 22.8%; Pred. No. 0.00029;

Matches 103; Conservative 57; Mismatches 172; Indels 119; Gaps 18;

QY 24 IALTFVGL--LGMMAFAPLDSAVI-ANGVSAEVSQVHLEGLMLAKILVREGEKVKAG 80
 DB 171 VILLIVGFCLGFARQTLYGAVVPADGMIALTTTPQ-----SGVYANVGVVQGRVAG 224
 QY 81 QVLFLDPTQANAAAGITRNOYVALKAMEARLAEERDQPSISFPADLTSGRADPVARA 140
 DB 225 QVLFVL-----AAE-----HHDDR-----GRPSQQAAY 248
 QY 141 IADDEQAQFTREROTIQQVDLNNARLQYQSEIEGIDROTGLKDLQGFIEDELIDRLK- 199
 DB 249 IAEQO-----RLTAEMAVQLRAQGLQOQAAA-----RALAGLRNLEQVDAELGVLRR 298
 QY 200 -----LYDKGLVPRPLALAEARAGSLSGISGRLTADRSKAVQASD 241.
 DB 299 QOLTFIEQRVYALTIRGLVSOQFVDEKQADVLDGRANL-----ELQRERLTLADALAQ 353
 QY 242 TQLKVRQIKQEFPEVOS-----QSTTERVRLAEVTEKEVVASDAQRIKIVSPVNGRA 295
 DB 354 AQAEIQQLPVSILRQDLALAGASLQADRRTAIDAAASREVRAPRGR----- 401
 QY 296 QNLRFTEGAVVRAEPLVDIAPEDEAFVIOAHFOPTDVNVHMGVTEVR--LPAFR- 351
 DB 402 VALRPLRGQAVGQGRDLPLPTSTATEVLYAPSRAGLIGPGLPVQRLRFDALPYGHY 461
 QY 352 -----SAGNPDERHDPVAVADRISDPQKARLFLGIYRVDKQLPHLRGRVTAG 402

THIS PAGE BLANK (USPTO)

Query Match	6.6%;	Score 142;	DB 10;	Length 355;
Best Local Similarity	21.2%;	Pred. No. 0.00082;		

Query Match	6.08;	Score 130;	DB 10;	Length 285;
Best Local Similarity	21.28;	Pred. No. 0.0057;		

Matches 87; Conservative 56; Mismatches 103; Indels 164; Gaps 20;

```

OY 23 IIALTFVGLLGMAFA-----PLDSAVIANGVSAEVSODVOHLESGMLAKILVREGGEYK 78
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 14 VAAAVLAGIMRYMLSPWTRDARADVVY---VADY---SGWTDLEVDKDNQYVK 66
OY 79 AGCVLFELDPQAAAAGITNNOYALKAMEARLLAEKDPSPISFPADLTQSRADPVA 138
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 67 VGDVLMRID-----OERYQANLEO-----A 86
OY 139 RAIDEQAQFERRQTIQGVDMNAQRLQYSEIEGIDROTQGLKDLGFIEDLIDLR 198
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 87 RAVAFETRQ-----QY-----LLRONEAR-----106
OY 199 KLYDKGLVPRRLALEARAGSLSGISGRLTADRSKAVOGASDTOLKVRQIKOFEFEQVS 258
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 107 -----RSRL-----GIGASIAEDKE-----MNOINAIARSEYQELA 139
OY 259 QSITEIRVRLAEVEKEVVASDAOKRIKIVSPVNGTAONLR-----FTEGAVVRAAEPL 313
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 140 Q-----VKIAELN-----LKRSELRAARGOVNLELAQGNVATAGQAVMA--L 181
OY 314 VDIAPDEAFIQAFOPTDVDMVMGMVTEVRLPAFHSAGNPDPERHDPVA--VADRIS 371
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 182 VD-----QOSFYVAYFEETKLPGRVGMRAQVRL---MSGDPIDGTYESISSGITTDRNS 234
OY 372 DPORQ-----ARLFLGI-VRVYDVKQLPPLRLGRVTAGMPAOYIV 409
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 235 TPDQQLANVEPTFNWVRLAQRIPIRIRLDQVADV--HLSAGMTASTVT 282

```

RESULT 7
US-09-820-843A-113
Sequence 113, Application US/09820843A
Publication No. US2003039963A1

```

GENERAL INFORMATION:
APPLICANT: Council of Scientific and Industrial Research
TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PROTEIN
FILE REFERENCE: 063915
CURRENT APPLICATION NUMBER: US/09/820, 843A
CURRENT FILING DATE: 2001-03-30
NUMBER OF SEQ ID NOS: 118
SOFTWARE: PatentIn version 3.0
SEQ ID NO 113
LENGTH: 2354
TYPE: PRT
ORGANISM: L. major
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: AC005802_5 L6202.3
NAME/KEY: misc_feature
OTHER INFORMATION: g116899670
US-09-820-843A-113

```

Query Match 5.98; Score 128.5; DB 9; Length 2354;
Best Local Similarity 25.3%; Pred. No. 0.15;
Matches 60; Conservative 31; Mismatches 97; Indels 49; Gaps 9;

```

OY 101 QYVALKAMEARLLAERDQ-RPSISFPADLTQSRAD-----PWVARIADE---QA 146
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 727 QRALEAGVARIADRDREARQOLANAELOQLDITATQRAELEAOULARIADRDREARQ 786
OY 147 QFERRQTIQGVDMNAQRLQYSEIEGIDROTQGLKDLGFIEDLIDLRKLYDKGLV 206
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 787 QLANAEELQORLDTATQRAELEAGVARIADRDREARQOLANAEEL-----834
OY 207 PRPPLALAEARAGSLSGISGRLTADRSKAVOGASDTOLKVRQIKOFEFEQVSITETRV 266
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 835 -QORLDTATQORALEAGVARIADRDREARQ-----OLAA-----NAEELQRLDTATQORA 885
OY 267 RL-AEYVEKEVVASDAOKRIKIVSPVNGTAONLRFTFEGAVVRAAEPLVDIAPDEA 322
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Db 886 ELEAGVARIANAELEQRL-----DTATQORALEARVARILA-----ADRDEA 929

```

RESULT 8
US-09-991-496-120
Sequence 120, Application US/09991496
Patent No. US20020169285A1
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Campos-Neto, Antonio
APPLICANT: Webb, John R.
APPLICANT: Dillon, Davin C.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Bhatia, Ajay
APPLICANT: Colier, Rhea
APPLICANT: Probst, Peter
APPLICANT: Brannon, Mark
TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
THERAPY AND DIAGNOSIS OF LEISHMANIASIS
FILE REFERENCE: 210121.420C9
CURRENT APPLICATION NUMBER: US/09/991.496
CURRENT FILING DATE: 2001-11-20
NUMBER OF SEQ ID NOS: 137
SOFTWARE: FastSeq for windows Version 4.0
SEQ ID NO 120
LENGTH: 2310
TYPE: PRT
ORGANISM: Leishmania major and chagasi
US-09-991-496-120

```

Query Match 5.98; Score 127; DB 9; Length 2310;
Best Local Similarity 23.4%; Pred. No. 0.2;
Matches 75; Conservative 46; Mismatches 128; Indels 72; Gaps 13;

```

OY 40 LDSAVIANGVSAEVSODVOHLESGMLAKILVREGGEYK-----AGVLFELDPQOANAA 94
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1318 LDTATQORALEARVAR--LAADGDEARQOLANAELEQRLDTATQORALEAGVARI 1375
OY 95 AG-----ITRNOYVALKAMEARLLAERDQ-RPSISFPADLTQSRAD-----134
Db 1376 ANAELEQRLDTATQORALEARVARILADRDREARQOLANAELEQRLDTATQORALE 1435
OY 135 PWVARIADE---QAFERRQTIQGVDMNAQRLQYSEIEGIDROTQGLKQGLFIE 191
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1436 AQVARILADRDREARQOLANAELEQRLDTATQORALEARVARILADGDEARQOLANA 1495
OY 192 DELIDRLKLYDKGLVPRRLALAEARAGSLSGISGRLTADRSKAVO--GASDTOLKVRQI 249
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1496 EEL-----QORLDTATQORALEAGVARIADRDREARQOLANAEEL-----1537
OY 250 KOFEFEQVSITETRVRLAEVT-----EKEVVASDA--OKRIKIVSPVNGTAONLRFF 301
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1538 -QORLDTATQORALEARVARILADGDEARQOLANAELEQRL-----DTATQORAE 1589
OY 302 TEGAVVRAAEPLVDIAPDEA 322
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1590 LEARVARIA-----ADRDEA 1604

```

```

RESULT 9
US-09-820-843A-114
Sequence 114, Application US/09820843A
Publication No. US2003039963A1
GENERAL INFORMATION:
APPLICANT: Council of Scientific and Industrial Research
TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PRO
FILE REFERENCE: 063915
CURRENT APPLICATION NUMBER: US/09/820, 843A
CURRENT FILING DATE: 2001-03-30
NUMBER OF SEQ ID NOS: 118
SOFTWARE: PatentIn version 3.0
SEQ ID NO 114

```


Db 218 VAOSDALKAAEQADAEIERLADFCGLNDRSMLNDVIGLLDERESLASAESELAQAR 277
QY 199 KLYDKGVLPRRLALEARAGSLSGSIGRLT----- 229
Db 278 AAGDLEAV-----AAAEKAVAGLEQSIASKTSTWPSQDQTYLQSYTALEAEARRYASTTE 332
QY 230 -----ADRSKAVOGASDQ-----LKVKOIKOEFEEVOVSOSITE 263
Db 333 ALEIAERLYIDSLGRVSELSAAQRAVAEASAOQDALGETAQLSTOHOLEAQSSAID 392
QY 264 TRVRLAEYTEREYVASDAQRIKI-----VSPVNGTAONLFFETEGAVVRAAEPLVDIAP 318
Db 393 AALGLASVDNKAATRSTQGLKMDINNTVRSPIYSIVSSVC-AAQGC--PAGALLSYA- 448
QY 319 EDEAFVIOAHFQPTDQVNVHGMWTEVRLPAFHSAGNDP---PERHDPVAVA 367
Db 449 DDSEIKITANVKEAISNVTIGSRVTFPTP---STGTKEFAGRVSKVPIAIA 498

RESULT 12

US-09-815-242-5064
; Sequence 5064, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5064
; LENGTH: 2472
; TYPE: PRF
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-5064

Query Match 5.3%; Score 115.5; DB 10; Length 2472;
Best Local Similarity 20.6%; Pred. No. 1.9;
Matches 112; Conservative 78; Mismatches 178; Indels 175; Gaps 25;

QY 37 FAPLDSAVIA-----NGVSAEVSQDVHLEGG-MLAKI-----LYRE----- 73
Db 1534 FDLIDSAAALQRMDDVDVNTIELEALQDRLHTLKGARMAIEIGDGLAHELFVEGL 1593
QY 74 -GEVVKAGGYLF-----ELDPTQANAAAGITRNOYVALKAMEARLLAERDQRPSTSF 124
Db 1594 CGGRLRSPALFGLLQCHDELAEMLAIVRGHRTLPGQALIAIRLRLRSDPDQLSVP- 1652
QY 125 PADLTSRAPPMVAVRAIADBOAQ-----FTERRQ----- 153

Db 1633 ----TSVSLKPLAKGAANADESEILIDFLEADLLENLELALGHWGNGDQAPLDDL 1708
QY 154 -----TTCGVDDLM-----AQR-----LOYOSLIGIDROTQ 182
Db 1709 RILHTLKCARLAAQGTETELGNLAHDECHLTDAQOGGAWPSPSLILDAGSGLEGQROYDL 1768
QY 183 LKDLGFTIEDELIDRLKLYDKGLVPRRLALEA---RAGS-----LSGSIGRLTA 230
Db 1769 LRERLA--EDD-----EAGEREPQALVQADDTDRAVASALAELETRLAPAGAIMA 1818
QY 231 DRSKAVOGASQOLKVKROIKOEFEVOVSOSTEIRVR-AYVTEKEV---VASDAQRI 285
Db 1819 AEAPPAAPATLTPVPRKQAQDAQASRRAPQDELKVPABULELVNLAGESTIFGRV 1878
QY 286 -KIVSPVNGTAONLFFETEGAVVRAAEPLVDIAEDEAFVIOAH-----FOPTD 333
Db 1879 EQVSDVQFTLGEH-----ESTIERVRODLRLDLETQALISRBQDAERAGYEEDFDE 1934
QY 334 VDNVHGMWTEVRLPAFHSAGNDPDERHDPVAVADRISDPQKARFL----- 381
Db 1935 MD--RYSOLOLSRALFESAS-----DLIDLKETLAKNRDDETLILQARVNTLQ 1985
QY 382 GIVR---VDVROLPPHLR---GRVTAGMPAOVIVPTG-----ERTVLOYFSPLRDTLR 429
Db 1986 GLMRTWVPFDRLVPRRLRRIYROYAVAGELGQVEFVYVNADEMDRTVLERIVAPLEHMLR 2045
QY 430 TTM 432
Db 2046 NAV 2048

RESULT 13

US-09-815-242-11146
; Sequence 11146, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11146
; LENGTH: 390
; TYPE: PRF
; ORGANISM: Haemophilus influenzae
US-09-815-242-11146

Query Match 5.3%; Score 114.5; DB 10; Length 390;
Best Local Similarity 20.4%; Pred. No. 0.16;

Matches 93; Conservative 65; Mismatches 138; Indels 159; Gaps 21;

```

OY 23 IATLEFVGL---LGNAAF-----APLDSAVIANGV-VSAEVSQDVQHLEGMIAKTIVRE 73
Db 28 ILLIILITIGACLYIFFELFKDEEETEDAVGNGVWSSQVA-----GNVAKIMADN 79
OY 74 GERKAGOVLELDPDQANMAAGITRNOYVALKAMEARLLAERDORPSISFPADLTSGRA 133
Db 80 MDKVAAGDVLVELDPTNKK-----LSFEBA 104
OY 134 DPMVABALAD-EQAOFTERROTIQGOVDLMAORLOYSSEIGIDROTGKLDQGLTFED 192
Db 105 KSNLNAAROVQOGLFTV--OOLOSVAHANEISLQAOGNLA---RRVQ--LEKMGALDK 157
OY 193 ELIDLRKLYDKGLVRRPRLALAEARAGSLSGISGRITAD---RSKAVGASDGTQKLVQ 248
Db 158 ESFOHAK-----EAVELAKANLAKSKNOLANQALLRNVLPEOPQIOANLS 205
OY 249 IKOEFEQVSOSITETRVRLAEVTEKEVVASDAOKRIKIVSPVNGTAQMLREFTEGAVVR 308
Db 206 LKQANLNL-----LQRTKRSPIDGIVAR-RNVQVQANVS 238
OY 309 AAEPLVDIAPEDEAVTQAHFOPTDVNVHGMVTEVR----- 346
Db 239 VCGALMAVVSNOQM-LEANFKETOLTNRIGQPVKIHFDLYGKNKEFGDVGINGEMGTG 297
OY 347 -----LPAFHSAGNPDPREHDPVAVADRIS-----DPOK--QARLELGIVRDVQOLPRH 394
Db 298 NAFSLPSQONATGN-----WIKVQVRVPRIKIDPOOFTETPLRIG-----LSAT 342
OY 395 LRGVYTAGMPAQVIVPTGERTVLOYLFSPLRDTLR 429
Db 343 AKVRISDSGAML-----REKTEPKTLFS--TDTLK 371

RESULT 14
US-09-883-343A-6
; Sequence 6, Application US/09883343A
; Publication No. US20030039632A1
; GENERAL INFORMATION:
; APPLICANT: Stiles, Michael E.
; APPLICANT: Vederas, John C.
; APPLICANT: van Belkum, Marius J.
; APPLICANT: Morobo, Randy W.
; APPLICANT: Greer, G. Gordon
; APPLICANT: McMullen, Lynn M.
; APPLICANT: Lelsner, Jorgen J.
; APPLICANT: Poon, Alison
; APPLICANT: Franz, Charles M.A.P.
; TITLE OF INVENTION: No. US20030039632A1eBacteriocins, Transport and Vector System at
; FILE REFERENCE: 660,000505
; CURRENT APPLICATION NUMBER: US/09/883,343A
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US/08/924,629
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 60/026,257
; PRIOR FILING DATE: 1996-09-05
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 457
; TYPE: PRT
; ORGANISM: Lcadi;
US-09-883-343A-6

Query Match 5.2%; Score 113.5; DB 9; Length 457;
Best Local Similarity 18.4%; Pred. No. 0.25;
Matches 85; Conservative 76; Mismatches 195; Indels 105; Gaps 16;
OY 7 QREPTNFOAVARIGIITLFTVGLGMAFAPLDSAVIANG-VSAEVSQDVQHLEGM 65
Db 14 QRRYRNFPLIIVP--IFLAVVFILFSLFAKREIIVKASGEIIPAKVLSIDQSTNNA 70

```

```

OY 66 LAKIIVREGKVKACOVLELDPDQANMAAGITRNOYVALKAMEARLLAERDORPSI--- 122
Db 71 IDSNOLTEKNKVKKGDVLTFTTSGNEKISSQLTQO---LNNLNDRLKSLDITTKOSIVNG 127
OY 123 -----SFPAD-----LTSQADPMVABALADQAOFTERROTIQGOVDLM 162
Db 128 RSEFGGTDOFGYDSLFGNGMAQVDLTSEFNQOSSDKOTADQOANH-----QIDVL 178
OY 163 -----ANQRLQYSEIGIDROTGKLDQGLTFED---ELIDLRKLYDKGLVRRPL 211
Db 179 KQGSKNNOQLANYQAILTSINSNTRKPTNNPYQATIDNSADLKSAOTTDGDDQVAKOTAL 238
OY 212 L-----ALEARAGSLSGISGRITFDRSKAVGASDGTQKROIKOEFEQVSOSITET 264
Db 239 SNVQOQIDLOLTSSSYDSQIAGIT---KSGPLSOSTIDKTLADLKQQLASAKREINDQ 295
OY 265 RVRLAEVTEKEVVASDAOKRIKIVSPVNGTAO-----NLREFTEGAVVRAEPLVDIAP 318
Db 296 QOSLDELKAKQSSANEDYQDVIAKPEDGILHATDKTKIKYEPKGTITAIQIYPKL---T 352
OY 319 EDEAVTQAHFOPTDV-----DNVHGMVTEVRLPAFHSAGNPDPREHDPV 364
Db 353 QRTALNVEEYVPASNITGLKQRAIRFVANQNVTRPLTLNGTIKSISSA-----PI 403
OY 365 AVADRISDPQKARLF-----LGIYRVQKOLPRHLRGHTV 400
Db 404 A-----SKESFYKLVATTQAKIDNEQIKYGLNGRIT 436

```

RESULT 15
US-09-815-242-5724

```

; Sequence 5724, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5724
; LENGTH: 792
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(792)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-815-242-5724

```

